On the index of multi-mode DAE Systems (also called Hybrid DAE Systems)

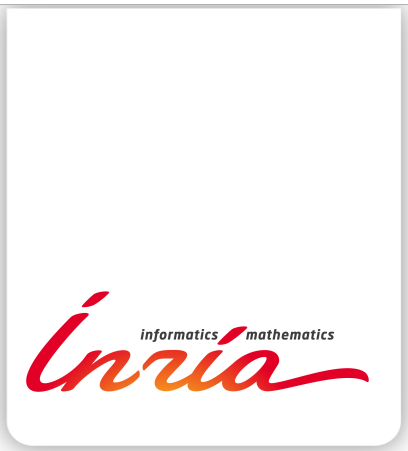
Albert Benveniste, Timothy Bourke, Benoît Caillaud, Marc Pouzet

On the index of multi-mode DAE systems (also called Hybrid DAE systems)

Albert Benveniste, Timothy Bourke, Benoît Caillaud, Marc Pouzet

RESEARCH REPORT
N° 8630
November 2014
Project-Teams Hycomes, Parkas

ISSN 0249-6399 ISRN INRIA/RR--8630--FR+ENG



On the index of multi-mode DAE systems (also called Hybrid DAE systems)

Albert Benveniste*, Timothy Bourke†, Benoît Caillaud‡, Marc Pouzet§

Project-Teams Hycomes, Parkas

Research Report n° 8630 — November 2014 — 30 pages

This work was supported by the SYNCHRONICS "action d'envergure" of Inria and by ITEA/Modrio project.

* INRIA, Rennes, France. corresp. author: Albert.Benveniste@inria.fr
† Ecole Normale Supérieure (ENS), Paris
‡ INRIA, Rennes, France
§ Ecole Normale Supérieure (ENS), Paris

RESEARCH CENTRE
RENNES – BRETAGNE ATLANTIQUE

Campus universitaire de Beaulieu
35042 Rennes Cedex

Abstract: Hybrid systems modelers exhibit a number of difficulties related to the mix of continuous and discrete dynamics and sensitivity to the discretization scheme. Modular modeling, where subsystems models can be simply assembled with no rework, calls for using Differential Algebraic Equations (DAE). In turn, DAE are strictly more difficult than ODE.[¶] In most modeling and simulation tools, before simulation can occur, sophisticated pre-processing is applied to DAE systems based on the notion of *differentiation index*. Graph based algorithms such as the one originally proposed by Pantelides are efficient at finding the index, structurally (i.e., outside some exceptional values for the system parameters). The differentiation index for DAE explicitly relies on everything being differentiable. Therefore, extensions to hybrid systems must be done with caution—to our knowledge, no such extension exists.
We propose to rely on *non-standard analysis* for this. Non-standard analysis formalizes differential equations as discrete step transition systems with infinitesimal time basis. We can thus bring hybrid DAE systems to their non-standard form, where the notion of *difference index* can be firmly used—the difference index of a difference Algebraic Equation (dAE) is an easy transposition of the differentiation index, in which forward shift replaces differentiation. We prove that the differentiation index of a DAE is structurally equal to the difference index of its non-standard interpretation, which is a dAE. We can thus propose the difference index of the non-standard semantics of a hybrid DAE system, as a consistent extension of both the differentiation index of DAE and the difference index of dAE.
It turns out that the index theory for (discrete time) dAE systems is interesting in itself and raises new issues. We discuss graph based algorithms à la Pantelides for computing the dAE index and discuss examples.

Key-words: Hybrid systems, DAE, index, nonstandard analysis

Sur l'index des DAE multi-mode (ou DAE hybrides)

Résumé : Les outils de modélisation de systèmes hybrides posent des difficultés liées au mélange continu/discret et à la sensibilité aux schémas de discrétisation. Pour que la modélisation soit modulaire, c'est-à-dire que le modèle global soit obtenu par simple assemblage de sous-modèles sans rien d'autre à faire, il faut avoir les DAE (Equations différentielles algébriques) qui sont des contraintes. Dans la plupart des outils de modélisation acceptant les DAE, une phase préliminaire existe qui consiste en le calcul de *l'index de différentiation.* Une famille d'algorithmes à base de graphes ont été proposés, d'abord par Pantelides, puis par plusieurs auteurs. Ces algorithmes calculent de manière efficace une valeur structurelle pour l'index (valide pour "presque toutes les valeurs" des coefficients).

La notion d'index de différentiation repose explicitement sur la différentiabilité. Par conséquent, l'extension de cette notion aux systèmes hybrides doit être conduite avec précaution. A notre connaissance, il n'existe pas de tel développement.

Nous proposons de nous appuyer sur *l'analyse non standard* à cet effet, car les équations différentielles y sont vues comme des équations aux différences (à temps discret) où le pas de temps est *infinitésimal*. Avec cette interprétation des DAE, on peut les regarder comme des systèmes à temps discret. Pour de tels systèmes on peut proposer une contrepartie de la notion d'index de différentiation, nous l'appelons *l'index aux différences*. Nous montrons que les deux notions d'index que l'on peut, ainsi, attribuer aux DAE coincident. Ceci nous permet de proposer, pour les systèmes hybrides à DAE, de définir leur index comme étant celui de leur interprétation nonstandard. Cette définition donne bien une extension conservative de l'index des DAE et de l'index de systèmes à temps discret.

Il se trouve que la notion d'index pour les systèmes dynamiques à temps discret est intéressante en soi. On propose des algorithmes à la Pantelides et nous décortiquons des exemples.

Mots-clés : Systèmes hybrides, DAE, index, analyse non standard

CONTENTS

I. Introduction

The booming of system design during the last decade has called for drastic changes in techniques and tools used for physical system modeling. So far the dominant technology for physical system modeling has been and still is Simulink,[1] [2] in which both discrete and continuous time systems must be expressed in state space form—we give here the continuous time version of it:

$$\left\{\begin{array}{lcl} \dot{x} & = & f(x,u) \\ y & = & g(x,u) \end{array}\right. \tag{1}$$

where u is the input vector, x the state vector, and y the output vector. Systems of the form (1) consist of Ordinary Differential Equations (ODE). They compose as input/output functions, which requires that no integrator-free loop is created as the result of the composition. Unfortunately, commonly encountered physical systems are naturally modeled by balance equations (Ohm and Kirchhoff laws for electrical networks, Lagrange equations for mechanical systems, thermal systems), which naturally leads to integrator-free loops. Manually moving from balance equations to state-space forms like (1) quickly becomes a significant burden. Since such a translation, from balance equations to state-space form, must be performed globally for the entire system, the reuse of partial models from libraries gets impaired [18], [20], [21].

The need for an adapted modeling framework was already recognized in the 70's by the engineering community with the proposal of *bond graphs* [13], [12], [16]. Bond graphs guide the development of models from first physical principles by identifying generic concepts such as power, effort, and flow. The *causality analysis* of a bond graph allows deducing who in the model should be considered as an input, state, or output. Model reuse occurs at the level of bond graphs, whereas the causality analysis is performed automatically and globally for the system, prior to simulation. The causality analysis of bond graphs is a graphical problem.

The mathematical counterpart of models from physical principles is that of a Differential Algebraic Equation (DAE) [8], [22], of which a simple instance is—again in continuous time:

$$\left\{\begin{array}{lcl} 0 & = & f(\dot{x},x,u) \\ 0 & \leq & g(x,u) \end{array}\right.$$

Here we have shown a combination of equality and inequality (also called bilateral and unilateral) constraints; x is the state vector collecting all variables subject to differentiation, whereas u is the algebraic vector collecting other variables. Observe that we do not mention inputs nor outputs. Unlike state-space systems, DAE systems compose with no restriction. Modeling languages based on DAEs were initially proposed in the early 80's by H. Elmqvist [11] from the automatic control group lead by K.J. Aström in Lund. The Modelica language [18], [20], [21] subsequently resulted from this effort, grounding the birth of the Modelica community.[3]

[1] www.mathworks.com/products/simulink/
[2] http://www.mathworks.com/help/simulink/
[3] https://www.modelica.org/

There is, however, no free lunch, and L.R. Petzold, one of the mathematicians having grounded the foundations of DAE theory, entitled one of his papers: "DAEs are *not* ODEs" [22]. As an example, consider the following system, which we state in both continuous time (left handside) and discrete time (right handside) versions:

$$\left\{\begin{array}{lcl} \dot{x} & = & f(x,u) \\ 0 & = & g(x) \end{array}\right. \qquad \left\{\begin{array}{lcl} x^\bullet & = & f(x,u) \\ 0 & = & g(x) \end{array}\right. \tag{2}$$

In (2) both x and u are scalar signals, taking their values in $\mathbb{R}$. For the discrete time version, $x^\bullet$ denotes the forward shifted version of signal x, i.e., such that $x^\bullet_k = x_{k+1}$ for every discrete instant $k = 0, 1, 2, \ldots$ Systems of the form (2) exhibit so-called *latent constraints*. For the discrete time version, additional constraint $0 = g(x^\bullet)$ follows by shift invariance, which furthers constrains the triple $(x^\bullet, x, u)$. Similarly, for the continuous time version, differentiating the second equation yields the additional constraint $0 = \frac{d}{dt}g(x) = g'(x)\dot{x} = g'(x)f(x,u)$ where g' denotes the derivative of g, which furthers constrains the triple $(\dot{x}, x, u)$. To summarize, (2) should rather be reformulated by making the latent constraints explicit (shown in red):

$$\left\{\begin{array}{lcl} \dot{x} & = & f(x,u) \\ 0 & = & g(x) \\ 0 & = & g'(x)\dot{x} \end{array}\right. \qquad \left\{\begin{array}{lcl} x^\bullet & = & f(x,u) \\ 0 & = & g(x) \\ 0 & = & g(x^\bullet) \end{array}\right. \tag{3}$$

Further shifting the last equation in the discrete time version of (3) would bring the fresh variable $x^{\bullet\bullet}$, denoted by $x^{\bullet 2}$, thus resulting in no further constraint on the triple $(x^\bullet, x, u)$. Similarly, further differentiating the last equation of the continuous time version of (3) does not bring anything useful. Focus once more on the discrete time version of (3), which we rewrite by substituting $f(x,u)$ for $x^\bullet$ in its last equation:

$$\left\{\begin{array}{lcl} x^\bullet & = & f(x,u) \\ 0 & = & g(x) \\ 0 & = & g(f(x,u)) \end{array}\right. \tag{4}$$

Assume the current value of state x is *consistent*, meaning that constraint $g(x) = 0$ is satisfied. Then, the last equation determines u as an output of the system and then, the pair (x,u) determines the next value for the state $x^\bullet$ by using the first equation, ensuring that $x^\bullet$ remains consistent. This was an instance of the kind of *causality analysis* we can perform to infer the input/output status of the different algebraic variables. The same analysis can be performed for the continuous time version of (3), replacing forward shift by differentiation. One point should be noticed: it could happen that the last two equations of (4) are redundant. This would occur if $f(x,u) \equiv x$. In this case, our causality analysis can still be performed but leads to incorrect conclusions: u is not an output any more as it disappears from the equations of the system. The kind of causality analysis that we performed leads to valid conclusions but for exceptional instances of the functions f and g: properties inferred in this way are called *structural*.

The process of revealing latent constraints by shifting or differentiating is called *index reduction* and the number of

times it has to be done is called the *differentiation* or *difference index,* for continuous time and discrete time systems, respectively. The notion of differentiation index was proposed in the late 80's, see [7], [24]. Index reduction goes along with causality analysis, performed by graph based algorithms proposed for the first time by Pantelides [19] and subsequently developed and modified by several authors [25]. Some physical modeling tools use index reduction to prepare for the work of the solvers, which are designed to simulate DAE of low index only $(1, 2,$ or at most 3). Some solvers address the issue of latent constraints (including invariants such as occurring in Hamiltonian or Symplectic systems) in an implicit way, by using appropriate discretization schemes [2]. Nevertheless, index reduction was originally proposed to find consistent initial conditions for the DAE and still remains useful for that purpose. In addition, preprocessing required by some solvers for so-called non-smooth systems [1] amount to computing the *relative degree*, which appears closely related to the index.

Surprisingly enough, no notion of index exists for hybrid extensions of DAE. When equations are used in both the continuous dynamics (in each mode) and the reset of states at mode transitions, it becomes unclear how such resets must be performed, thus leading to latest physical systems modelers to refuse certain, otherwise mathematically well defined, models.

The intuition conjectures that the right notion of index should imply that, in each mode, the index should coincide with the known notion of index for DAE. This looks fine also when transitions between modes are isolated. How to handle cascades of mode transitions, however, seems out of reach of the intuition.

In this work, we show that the above intuition is indeed incorrect and we provide a formal definition of the index, for hybrid DAE systems. *While the notion of index remains global and "mode independent", the dependence on modes actually appears while performing the graph based causality analysis supporting index evaluation.* Our approach is as follows:

1) We observe that a notion of *difference index* can be defined for discrete time algebraic dynamical systems, we call them dAE, (i.e., involving constraints) by just borrowing, to discrete time dynamics, the notion of differentiation index.
2) We use the *nonstandard analysis* interpretation of DAE—a nonstandard interpretation of a DAE consists in interpreting the derivative as a difference operator using an "infinitesimal" step. This interpretation is exact in a certain sense, it is not an approximation unlike discretization schemes—it is not effective, however (there is no free lunch). We show that the two notions of index for a DAE (the classical differentiation index and the nonstandard difference index) coincide.
3) We take the nonstandard interpretation of a hybrid DAE system and consider its difference index. Thanks to the previous item, this notion subsumes the differentiation index of DAE and the difference index of dAE.

In fact, our study does not address the "true" index [24] (see the forthcoming Definition 2), but rather the *structural* index. For a linear DAE system, the index equals the structural index almost everywhere when the non-zero coefficients vary over a neighborhood, and this extends to nonlinear DAE systems by considering the tangent linear DAE. Highly efficient graph based algorithms exist for computing the structural index, including the well-known first one proposed by Pantelides [19]. This graph-based analysis proposed by Pantelides is indeed a causality analysis, very much related to the causality analysis of bond graphs.

As our study concerns the structural notions of index, the term index *will refer to the structural index in the sequel, whereas we use the term* true index *to refer to the original notion of differentiation index.*

The paper is organized as follows. The background on differentiation index theory for DAE is recalled in Section II. After recalling what the index is following the original definition given in Campbell and Gear [24], we introduce the structural index and relate it to the notion of structurally nonsingular matrices. Structural properties of matrices are characterized by the existence of certain permutation matrices, whose quest relies on graph based algorithms. Section III introduces the index theory for difference Algebraic Systems (dAE), the discrete time counterpart of DAE; dAE can also be seen as discrete time dynamical systems subject to a transition relation (not function); we focus on dAE for which the transition relation is specified as a set of equations. While the definition of the (true) index is a straightforward translation from the continuous time case, new issues arise, due to the consideration of nonsmooth systems. We devote subsection III-E to the discussion of illustration examples, thus motivating the new techniques of guarded causality analysis we develop in Section IV—these techniques turn out to be closely reminiscent of those used in the compilation of synchronous languages [5], [3]. In Section V we recall how DAE can be given a discrete time (dAE) interpretation by using nonstandard analysis with infinitesimals, which yields an alternative way of defining the index, based on this dAE interpretation. We show that, for smooth DAE systems, the two notions of structural index coincide, which allows us to rely on nonstandard semantics for the index theory of hybrid DAE systems. The latter is developed in Section VI. We propose a mini-formalism of guarded equations that is expressive enough to cover practical examples, including inequality constraints, also called "unilateral", and variations thereof. We give the nonstandard semantics of such systems, which provides as a byproduct a notion of structural index together with the graph based algorithm for computing it. Examples are revisited in Section VII and hints for effective modular algorithms for performing causality analyses and computing the index are sketched in Section VIII.

II. Background on differentiation index theory

In this section we recall the background on differentiation index theory. We pay particular attention to the foundations of graph based algorithms associated to the quest for the

structural index. To this end, we recall the background on structurally nonsingular matrices with some detail and we give proofs.

A. DAE differentiation index

The basic reference is the work of Campbell and Gear [24], see also Mattsson and Söderlin [25]. Let $\mathbb{R}$ denote the set of reals, $\mathbb{Z} = \{\dots, -1, 0, +1, \dots\}$ the set of integers, and $\mathbb{N} = \{0, +1, \dots\}$ the set of non-negative integers. In this section we consider time invariant DAE problems of the following form:[4]

$$F(x, \dot{x}) \quad = \quad 0 \tag{5}$$

where x takes its values in $\mathbb{R}^n$ and F in $\mathbb{R}^m$. In the sequel, F_x and $F_{\dot{x}}$ denote the partial derivatives of F with respect to the first and second variables of F, respectively. For the following definition, the reader is referred to [24].

Definition 1: DAE (5) is solvable[5] *in the connected open set $\Omega \subset \mathbb{R}^{2n}$ if there are connected open sets $\Lambda \subset \mathbb{R}^\rho$ and $\mathcal{I} \subset \mathbb{R}$ and a function $(t, \lambda) \to \Phi(t, \lambda)$ such that:*

1) *$\Theta(t, \lambda) = (t, \Phi(t, \lambda))$ is a diffeomorphism of $\mathcal{I} \times \Lambda$ into $\mathbb{R}^{n+1}$.*
2) *$\Phi(t, \lambda)$ is a solution of (5) for each value of λ.*
3) *$(\Phi(t, \lambda), \frac{d}{dt}\Phi(t, \lambda)) \in \Omega$ for every $\lambda \in \Lambda$ and $t \in \mathcal{I}$.*
4) *If $x(t)$ is a solution of (5) such that $(x(t), \dot{x}(t)) \in \Omega$ for some $t \in \mathcal{I}$, then it holds that $x(t) = \Phi(t, \lambda)$ for some $\lambda \in \Lambda$. A pair (t, x) is called* consistent *if $x = \Phi(t, \lambda)$ holds for some λ.*

Condition 2) expresses that λ acts as a daemon solving the possible nondeterminism. Condition 4) states that λ captures all the nondeterminism. Indeed, λ parameterizes consistent initial conditions, which, in turn, determine solutions of (5). Systems with exogeneous inputs, of the form, e.g.:

$$F(x, \dot{x}, u) \quad = \quad 0 \tag{6}$$

are used in control and when composing subsystems to form larger systems. Systems of the form (6) are a specialization of (5) by putting $y = (x, u)$ and reformulating it as a DAE with larger state y. Systems of the form (6) leave generally some freedom on exogeneous u (subject to the constraints) when selecting solution $\Phi(t, \lambda)$.

The true differentiation index for DAE (5) is defined as follows. The *kth derivative array* associated to (5) is:

$$\begin{bmatrix} F(x, \dot{x}) \\ \frac{d}{dt}F(x, \dot{x}) \\ \vdots \\ \frac{d^k}{dt^k}F(x, \dot{x}) \end{bmatrix} \quad =_{\text{def}} \quad \begin{bmatrix} F^{(0)}(x, \dot{x}, w) \\ F^{(1)}(x, \dot{x}, w) \\ \vdots \\ F^{(k)}(x, \dot{x}, w) \end{bmatrix}$$

$$=_{\text{def}} \quad F_k(x, v, w) \tag{7}$$

$$\text{where } v \quad =_{\text{def}} \quad \dot{x} \tag{8}$$

$$\text{and } w \quad =_{\text{def}} \quad (x^{(2)}, \dots, x^{(k+1)}) \tag{9}$$

where we recall that $\frac{d}{dt}F(x, \dot{x}) = F_x(x, \dot{x})\dot{x} + F_{\dot{x}}(x, \dot{x})\ddot{x}$, and so on for higher degree derivatives. In (9) $x^{(1)} = \dot{x}, x^{(2)} = \ddot{x}, x^{(3)} = \dots$ denote the successive derivatives of x.

The reason for considering the kth derivative array equations is the following. The additional equations $\frac{d^j}{dt^j}F(x, \dot{x}) = 0$ added when forming the array are implied by the original DAE system, due to its time-invariance. These additional equations, however, add new equations and new variables, namely some components of $x^{(2)}$ involved in these new equations. Some of the new equations may not bring fresh variables, but only reuse previous variables, which they further constrain. In this case, *latent constraints* get revealed. Such latent constraints were not "visible" in the original formulation of the DAE system but are a consequence of its time-invariance.

Following again [24], a value x is called *consistent* for (7) if there exists (v, w) such that

$$F_k(x, v, w) \quad = \quad 0 \tag{10}$$

seen as an algebraic equation. Given a consistent value x for (7), algebraic equation (10) will generally have a set of solutions for (v, w).

Definition 2: Assume that DAE (5) is solvable. The true differentiation index *of this DAE, denoted by ν_D, is the smallest index k such that v is uniquely determined by the algebraic equation (10) for any consistent value x for (7).*

That is, for $k \geq \nu_D$, the map

$$x \quad \to \quad \exists w.F_k(x, v, w) = 0 \tag{11}$$

defines v as a deterministic function of x. Since (7) is equivalent to the original DAE, (11) determines $\dot{x}$ and is, therefore, sort of an ODE that can be solved for x. At this point, it is worth detailing how quantification over time applies to the map defined by (11):

$$\forall t \in \mathbb{R} \quad : \quad x_t \quad \to \quad \exists w_t.F_k(x_t, v_t, w_t) = 0 \tag{12}$$

By the implicit function theorem, (11) is, locally around a triple (v_o, w_o, x_o), equivalent to the following condition:

$$\begin{array}{l} \text{the map } \; x \to \exists w.Av + Cw + Ex = 0 \\ \text{defines } v \text{ as a deterministic function of } x \end{array} \tag{13}$$

where A, C, E are the Jacobians

$$A = (F_k)'_v, \quad C = (F_k)'_w, \quad E = (F_k)'_x,$$

at the considered triple (v_o, w_o, x_o).

As an illustration example consider the well known pendulum example in Cartesian coordinates:

$$\begin{array}{rcl} \ddot{x} & = & Tx \\ \ddot{y} & = & Ty - g \\ L^2 & = & x^2 + y^2 \end{array} \tag{14}$$

In this example, $x, \dot{x}, y, \dot{y}$ are the states and T is an algebraic variable. The form (5) for DAE system (14) is:

$$\begin{array}{rcl} \dot{x} & = & u \\ \dot{u} & = & Tx \\ \dot{y} & = & v \\ \dot{v} & = & Ty - g \\ L^2 & = & x^2 + y^2 \end{array} \tag{15}$$

[4] We may also consider $F(t, x, \dot{x}) = 0$, but dependence on time t can always be removed by making t an additional variable obeying $\dot{t} = 1$.

[5] The term used in [24] is *geometrically solvable.*

This is not of index 0 since the Jacobian with respect to $\dot{x}, \dot{u}, \dot{y}, \dot{v}, T$ is singular, leaving T undefined since the last row is identically zero:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -x \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -y \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

So we must differentiate the system. Differentiating, in (15), the last equation twice yields:

$$\begin{array}{rcll} \dot{x} & = & u & (i1) \\ \dot{u} & = & Tx & (i2) \\ \dot{y} & = & v & (ii1) \\ \dot{v} & = & Ty - g & (ii2) \\ L^2 & = & x^2 + y^2 & (iii) \\ 0 & = & \dot{x}x + \dot{y}y & (iv) \\ 0 & = & \dot{u}x + \dot{x}^2 + \dot{y}^2 + \dot{v}y & (v) \end{array} \tag{16}$$

Unknowns of highest derivative order are $\dot{x}, \dot{u}, \dot{y}, \dot{v}, T$. Rewriting all equations $(i$–$v)$ in the form $0 = \ldots$ yields the following Jacobian for the equations involving $\dot{x}, \dot{u}, \dot{y}, \dot{v}, T$:

$$\begin{bmatrix} \frac{\partial(i1)}{\partial \dot{x}} & \frac{\partial(i1)}{\partial \dot{u}} & \frac{\partial(i1)}{\partial \dot{y}} & \frac{\partial(i1)}{\partial \dot{v}} & \frac{\partial(i1)}{\partial T} \\ \frac{\partial(i2)}{\partial \dot{x}} & \frac{\partial(i2)}{\partial \dot{u}} & \frac{\partial(i2)}{\partial \dot{y}} & \frac{\partial(i2)}{\partial \dot{v}} & \frac{\partial(i2)}{\partial T} \\ \frac{\partial(ii1)}{\partial \dot{x}} & \frac{\partial(ii1)}{\partial \dot{u}} & \frac{\partial(ii1)}{\partial \dot{y}} & \frac{\partial(ii1)}{\partial \dot{v}} & \frac{\partial(ii1)}{\partial T} \\ \frac{\partial(ii2)}{\partial \dot{x}} & \frac{\partial(ii2)}{\partial \dot{u}} & \frac{\partial(ii2)}{\partial \dot{y}} & \frac{\partial(ii2)}{\partial \dot{v}} & \frac{\partial(ii2)}{\partial T} \\ \frac{\partial(iv)}{\partial \dot{x}} & \frac{\partial(iv)}{\partial \dot{u}} & \frac{\partial(iv)}{\partial \dot{y}} & \frac{\partial(iv)}{\partial \dot{v}} & \frac{\partial(iv)}{\partial T} \\ \frac{\partial(v)}{\partial \dot{x}} & \frac{\partial(v)}{\partial \dot{u}} & \frac{\partial(v)}{\partial \dot{y}} & \frac{\partial(v)}{\partial \dot{v}} & \frac{\partial(v)}{\partial T} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -x \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -y \\ x & 0 & y & 0 & 0 \\ 2x & x & 2y & y & 0 \end{bmatrix}$$

which, by reordering the rows, yields the Jacobian:

$$\begin{bmatrix} x & 0 & y & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -x \\ 0 & 0 & 1 & 0 & 0 \\ 2x & x & 2y & y & 0 \\ 0 & 0 & 0 & 1 & -y \end{bmatrix} \tag{17}$$

Removing the first (red) row yields, under the condition $y \neq 0$, a *structurally nonsingular* Jacobian, meaning that it is nonsingular but for exceptional values for the pair (x, y). Hence, $\dot{x}, \dot{u}, \dot{y}, \dot{v}, T$ is determined as a function of other variables. When y is zero or close to zero, then, due to constraint (iii), x is not small and we can exchange the roles of x and y. Hence, the index was found equal to 2.

In principle, Definition 2 requires that we not only differentiate (iii) twice, but also $(i1$–$ii2)$. This would, however, introduce fresh variables $x^{(2)}, x^{(3)}, u^{(2)}, u^{(3)}, y^{(2)}, y^{(3)}, v^{(2)}, v^{(3)}$, which enter the w of (11); eliminating this w is simply achieved by ignoring the differentiation of $(i1$–$ii2)$.

Following (10), a consistent value for the tuple x, u, y, v must satisfy the following equations, obtained by substituting $\dot{x}$ using $(i1)$ and $\dot{y}$ using $(ii1)$ in (iv):

$$\begin{array}{rcll} L^2 & = & x^2 + y^2 & (iii) \\ 0 & = & ux + vy & (iv) \end{array} \tag{18}$$

The remaining equations form an ODE with highest order derivatives $\dot{x}, \dot{u}, \dot{y}, \dot{v}, T$, since the Jacobian is structurally nonsingular (outside a neighborhood of $y = 0$):

$$\begin{array}{rcll} \dot{x} & = & u & (i1) \\ \dot{u} & = & Tx & (i2) \\ \dot{y} & = & v & (ii1) \\ \dot{v} & = & Ty - g & (ii2) \\ 0 & = & \dot{u}x + u^2 + v^2 + \dot{v}y & (v) \end{array} \tag{19}$$

The reasoning regarding the Jacobian obtained by erasing the first row of matrix (17) refers to so-called *structural* properties of matrices, which we recall now.

B. Structurally nonsingular matrices

Structurally nonsingular matrices play a central role in finding the differentiation degree using graph based algorithms. We thus recall some basic material in this section.

Say that a property $P(x_1, \ldots, x_k)$ involving the real variables $x_1, \ldots, x_k$ holds *almost everywhere* if it holds for every $x_1, \ldots, x_k$ outside a subset of $\mathbb{R}^k$ having empty interior. Matrix A is called *structurally onto* if it remains almost everywhere onto (surjective) when its non-zero entries vary over some neighborhood. Square $n \times n$-matrix P is a *permutation matrix* if and only if there exists a permutation σ of the set $\{1, \ldots, n\}$ such that $p_{ij} = 1$ if $j = \sigma(i)$ and $p_{ij} = 0$ otherwise. Pre- and post-multiplication of a matrix A by a permutation matrix results in permuting the rows and columns of this matrix. The following result holds [**?**]:

Lemma 1: A is structurally onto if and only if there exist two permutation matrices P and Q, of appropriate dimensions, such that $PAQ = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$, where B_1 is square with nonzero diagonal.

Proof: See Appendix A1. ■

Lemma 1 specializes to the following simpler result:

Corollary 1: A is structurally nonsingular if and only if PA has a nonzero diagonal (all entries of the diagonal are nonzero) for some permutation matrix P.

Matrix A is structurally nonsingular if and only if the linear equation $Av = y$, where v is the unknown, has a unique solution for any y, for almost all values for the non-zero entries of matrix A. We will, however, be interested in the more general equation, to be solved for v:

$$\exists w \ : \ Av + Cw + x = \begin{bmatrix} A & C \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} + x = 0 \tag{20}$$

where x has dimension n, A is a $n \times p$-matrix, C is a $n \times q$-matrix, and v, w are of appropriate dimensions. In other words,

when considering equation $Av+Cw+x=0$, we see v as the unknown (or output), x as the input, and w as a don't care—we are not interested in the particular value for w, we just want it to exist. We formalize this next.

Definition 3: Say that the pair (A, C) is structurally nonsingular *if equation (20) uniquely defines v as a function of x when x is consistent, almost everywhere when the non zero entries of (A, C) vary over some neighborhood.*

The following result generalizes Corollary 1:

Lemma 2: Pair (A, C) is structurally nonsingular if and only if there exists a permutation matrix P, of dimension n, such that

$$PA = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} \quad \textit{and} \quad PC = \begin{bmatrix} C_1 \\ 0 \\ 0 \end{bmatrix} \tag{21}$$

where A_2 is square with nonzero diagonal and C_1 is structurally onto.

Proof: See Appendix A2. ∎

C. Graph based algorithms

In this section, we recall the well-known graph based algorithm, originally due to Pantelides [19], which implements the quest for pivoting, i.e., for the two permutation matrices A and C of Lemma 2—we specify the algorithm but otherwise pay no attention to its efficient implementation, unlike the original work of Pantelides.

We assume three disjoint sets $\mathcal{X}, \mathcal{V}, \mathcal{W}$ of underlying real-valued variables—we also indicate the generic notation for their elements:

$$x{\in}\mathcal{X}, v{\in}\mathcal{V}, w{\in}\mathcal{W}, \text{ and we set } \mathbb{X} =_{\text{def}} \mathcal{X} \uplus \mathcal{V} \uplus \mathcal{W}$$

where $\uplus$ denotes the disjoint union. For E a scalar equation, i.e., of the form $F(\dots) = 0$ where F is a real-valued linear function of variables belonging to $\mathbb{X}$, write

$$z \mathrel{\overline{\in}} E$$

to indicate that variable z occurs in equation E with a coefficient that is structurally non-zero. Let $S = \bigwedge_{i\in I} E_i$ be a system consisting of the conjunction of a set $\{E_i \mid i \in I\}$ of scalar equations over a finite subset of variables belonging to $\mathbb{X}$. For convenience, we shall identify S with its set $\{E_i \mid i \in I\}$ of constitutive equations. Define

$$x \mathrel{\overline{\in}} S \text{ if and only if } x \mathrel{\overline{\in}} E_i \text{ holds for some } i \in I$$

and denote by $\mathrel{\overline{\in}}_S$ the set of variables $x \in \mathbb{X}$ such that $x \mathrel{\overline{\in}} S$. By decomposing

$$\mathrel{\overline{\in}}_S \;=\; \underbrace{\mathrel{\overline{\in}}_S \cap \mathcal{X}}_{\mathcal{X}_S} \uplus \underbrace{\mathrel{\overline{\in}}_S \cap \mathcal{V}}_{\mathcal{V}_S} \uplus \underbrace{\mathrel{\overline{\in}}_S \cap \mathcal{W}}_{\mathcal{W}_S}$$

and collecting all variables $z \in \mathcal{X}_S$ into a vector x, all variables $z \in \mathcal{V}_S$ into a vector v, and all variables $z \in \mathcal{W}_S$ into a vector w, system S uniquely defines an equation of the form (20). In the sequel,

$$\frac{\text{Prop}_1}{\text{Prop}_2} \quad \text{stands for} \quad \text{Prop}_1 \textit{ entails } \text{Prop}_2 \tag{22}$$

Definition 4: To each system $S{=}\{E_i \mid i{\in}I\}$, we associate its Pantelides graph $\mathcal{P}_S$*, which is a nondirected bipartite graph defined as follows, using notation (22):*

- *Its set of vertices is $\mathrel{\overline{\in}}_S \uplus S$;*
- *For every $z \in \mathrel{\overline{\in}}_S$:*

$$\frac{E \in S \quad \text{and} \quad z \mathrel{\overline{\in}} E}{z{-}E \in \mathcal{P}_S} \tag{23}$$

Call consistent causality *of S any directed bipartite graph $\vec{\mathcal{P}}_S$ such that:*

1) *Its set of vertices is $\mathcal{Z} \uplus S$, for some $\mathcal{Z}$ such that $\mathcal{X}_S \uplus \mathcal{V}_S \subseteq \mathcal{Z} \subseteq \mathrel{\overline{\in}}_S$;*
2) *$\vec{\mathcal{P}}_S$ covers $\mathcal{P}_S$: for every $z \in \mathcal{Z}$, $z{-}E \in \mathcal{P}_S$ if and only if either $z{\to}E \in \vec{\mathcal{P}}_S$ or $E{\to}z \in \vec{\mathcal{P}}_S$;*
3) *$\vec{\mathcal{P}}_C$ has the* single assignment *property, meaning that, for every z, there exists at most one equation E such that $E{\to}z \in \vec{\mathcal{P}}_C$;*
4) *Graph $\vec{\mathcal{P}}_S$ is circuitfree—hence its transitive closure is a partial order $\preceq_S$ on $\mathcal{Z}{\uplus}S$, we call it the* causality order *induced by $\vec{\mathcal{P}}_S$;*
5) *$\mathcal{X}_S \supseteq \min(\preceq_S)$, the set of minimal elements of $\preceq_S$;*
6) *For every pair $(z, w) \in (\mathcal{X}_S \uplus \mathcal{V}_S){\times}(\mathcal{W}_S \cap \mathcal{Z})$, it never holds that $w \preceq z$.*

What does the existence of a consistent causality ensure? By condition 1), $\vec{\mathcal{P}}_S$ involves as vertices all variables belonging to $\mathcal{X}_S$ (the "candidate inputs"), all variables belonging to $\mathcal{V}_S$ (the "outputs"), plus some additional variables belonging to $\mathcal{Z}$. By conditions 2) and 4), $\vec{\mathcal{P}}_S$ amounts to selecting a consistent orientation for $\mathcal{P}_S$, making it circuitfree: this defines the pivoting order between variables and equations, since, by Condition 3, at most one equation defines each variable. Condition 5) states that inputs for the system are found within $\mathcal{X}_S$: since they are minimal for the order $\preceq_S$, they get evaluated first (when reading the values of the inputs). Finally, condition 6) expresses that no variable belonging to $\mathcal{X}_S{\uplus}\mathcal{V}_S$ requires the prior evaluation of a variable belonging to $\mathcal{W}_S$, for its own evaluation. Thus eliminating the w variables simply proceeds by discarding certain equations.

The above discussion is formalized by the following theorem, where we associate with S the equation of the form (20) it defines, together with its pair (A, C) of matrices:

Theorem 1: System S possesses a consistent causality if and only if the pair (A, C) is structurally nonsingular.

Proof: See Appendix A3. ∎

The above results apply to the nonlinear equation $\exists w.F(x, v, w) = 0$ by considering the Jacobian ∇F at a solution (x_o, v_o, w_o) of $F(x, v, w) = 0$.

III. Index of difference Algebraic Equations

In this section we translate the theory of differentiation index to discrete time systems consisting of difference Algebraic Equations (dAE), the discrete time counterpart of DAEs. The translation simply consists in replacing the derivative operator by the forward shift operator. While the definition of the

true index is a straightforward translation, getting graph based algorithms is much more involved, as we shall see.

A. Difference Algebraic Equations (dAE)

A difference Algebraic Equation (dAE) consists of a constraint relating tuples of variables x and $x^\bullet$:

$$(x, x^\bullet) \in C, \quad \text{where} \quad C \subseteq D \times D \tag{24}$$

meaning that D is the domain of both x and $x^\bullet$.

Definition 5: A solution *of dAE* (24) *is any infinite sequence* $\{x_k \mid k \in \mathbb{Z}\}$ *satisfying*

$$\forall k \in \mathbb{Z} \quad : \quad (x_k, x_{k+1}) \in C$$

and (24) *is* solvable *if solutions for it exist.*

Definition 5 makes dAE (24) shift-invariant and expresses that

$$x^\bullet \text{ is the } \textit{forward shifted} \text{ version of } x\text{: } x^\bullet_k = x_{k+1} \tag{25}$$

For convenience, we write in the sequel

$$C(x, x^\bullet) \tag{26}$$

instead of (24).

B. The true index of a dAE

Referring to (9), we define the *kth difference array equations* associated to (26), which collects forward shifted versions of constraint $C(x, x^\bullet)$:

$$\begin{aligned}
\begin{bmatrix} C(x,x^\bullet) \\ C^\bullet(x,x^\bullet) \\ \vdots \\ C^{\bullet k}(x,x^\bullet) \end{bmatrix} &=_{\text{def}} \begin{bmatrix} C^{(0)}(x,x^\bullet,w) \\ C^{(1)}(x,x^\bullet,w) \\ \vdots \\ C^{(k)}(x,x^\bullet,w) \end{bmatrix} \\
&=_{\text{def}} C_k(x,v,w) && (27)\\
\text{where } v &=_{\text{def}} x^\bullet && (28)\\
w &=_{\text{def}} (x^{\bullet 2}, \ldots, x^{\bullet k+1}) && (29)\\
\text{and } x^{\bullet k+1} &=_{\text{def}} (x^{\bullet k})^\bullet
\end{aligned}$$

A value x is called *consistent* for (27) if there exists (v, w) such that

$$C_k(x, v, w) \text{ holds}, \tag{30}$$

seen as an algebraic equation. Given a consistent value x for (27), algebraic equation (30) will generally have a set of solutions for (v, w). Writing w_k instead of w in (27), the chain of sets $\mathcal{V}_k =_{\text{def}} \{v \mid \exists w_k : C_k(x, v, w_k)\}$ is decreasing for set inclusion. Having finite index for the considered dAE means that this chain becomes a singleton for some finite value of k, and then remains so. The true difference index of a dAE system is defined with reference to the true differentiation index for DAE systems:

Definition 6: Assume that dAE (24) *is solvable. The* true difference index *of this dAE, denoted by* ν_d*, is the smallest index* k *such that* v *is uniquely determined by the algebraic equation* (30) *for any consistent value* x *for* (27)*.*

That is, for $k \geq \nu_d$, the map

$$x \quad \rightarrow \quad \exists w. C_k(x, v, w) \tag{31}$$

defines v as a deterministic function of x. Since (27) is equivalent to the original dAE, (31) determines $x^\bullet$ and is, therefore, an OdE (Ordinary difference Equation), i.e., a transition system that can be directly executed. In (31), quantification over time applies as in (12), namely, (31) expands as

$$\forall n \in \mathbb{N} \quad : \quad x_n \quad \rightarrow \quad \exists w_n. C_k(x_n, v_n, w_n) \tag{32}$$

Again, we are rather interested in the *structural difference index,* simply referred to as *difference index* in the sequel, which is the almost everywhere value of the true difference index when the equations have their parameters modified while preserving the involvement/non-involvement of the variables in the equations.

C. Constructive semantics

Suppose that C has > 0 index, i.e., is not an OdE (a transition system in the usual sense). Then C_k involves in its second block-row the constraint $C^\bullet$ and thus, condition (31) ensures that $x^\bullet$ is also consistent. That is, if the system has a finite index k, considering the array C_k ensures that, at each reaction, current values for the variables and next values for the states are selected in a way that no deadlock will be caused in the future. For discrete time systems, the index appears as the needed look ahead horizon to be taken into account when generating the successives transitions of the system. The execution scheme of C is then specified by the following *constructive semantics*:

Constructive Semantics 1:

1) *Initial condition:* find a *consistent* x_0, i.e., such that
$$\exists w, \exists v : C_k(x_0, v, w) \text{ holds.}$$
2) *Non terminating while loop:* for every $n \geq 0$,
   a) assuming x_n consistent, find the unique v_n solution for v of $\exists w : C_k(x_n, v, w)$ holds;
   b) set $x_{n+1} = v_n$, which is consistent by construction;
   c) repeat. □

Constructive Semantics 1 specifies how runs of C are effectively constructed while time progresses, assuming a (static) constraint solver at hand.

So far Execution scheme 1 is global, so that the entire burden is on the shoulders of the constraint solvers. In the next sections we show how calls for constraint solvers can be made "local" for dAE specified as systems of equations, by developing causality analyses related to the notion of structural difference index.

D. Index of smooth dAE

By the implicit function theorem, for the special case in which constraint $C_k(x, v, w)$ has the form $F_k(x, v, w) = 0$ for F_k smooth, (31) is, locally around a triple (v_o, w_o, x_o), equivalent to the following condition:

$$\begin{aligned}&\text{the map } \; x \rightarrow \exists w. Av + Cw + Ex = 0\\ &\text{defines } v \text{ as a deterministic function of } x\end{aligned} \tag{33}$$

where A, C, E are the Jacobians

$$A = (F_k)'_v, \quad C = (F_k)'_w, \quad E = (F_k)'_x,$$

at the considered triple (v_o, w_o, x_o). Hence Lemma 2 can be invoked for a structural checking of (33) based on causality analyses, from which the (structural) difference index and then the constructive semantics follow. We do not develop details for this since our interest is rather in non-smooth dAEs.

E. Index of non-smooth dAE: examples

For smooth systems, index definition and analysis reduces to the structural analysis of Jacobians. Graph based algorithms rely on the so-called "incidence graph" consisting of branches linking each equation to the variables involved in it. The algorithms search for a consistent direction for each branch, making the graph directed and circuitfree. Theorem 1 provided the justification of the graph based algorithm of Section II-C for index evaluation.

In our study of hybrid systems, we will need to consider dynamical systems with modes. A mode is characterized by a predicate over the system variables. A hybrid system has a smooth dynamics in each of its different modes—the reader is referred to Section VI for a detailed description of this. The consequence is that we cannot restrict ourselves to smooth systems. The following simple examples illustrate the new difficulties arising with the consideration of non-smooth systems.

1) Guarded equations: Consider the following system of static equations:

$$\begin{array}{ll} E_0: & b = [x > 0] \\ E_3: & \textbf{if } b \textbf{ then } F_1(u,v) = 0 \textbf{ else } F_2(u,v) = 0 \end{array} \tag{34}$$

and call E_1 and E_2 the equations $F_1(u, v)$=0 and $F_2(u, v)$=0, respectively. Observe that E_0 is a function whose output is b. Then, a possible consistent causality for dAE system (34) in the sense of Definition 4 consists of the following directed branches:

$$\begin{array}{l} x \rightarrow E_0 \rightarrow b \\ b \rightarrow E_3 \\ u \rightarrow E_3 \rightarrow v \end{array} \tag{35}$$

(Exchanging u and v in the last causality would do as well.) The last two directed branches shown in the third line of (35) are legitimate according to Definition 4.

Now, assume that equations E_1 and E_2 have indeed the form $E_1 : v{=}f_1(u)$ and $E_2 : u{=}f_2(v)$, where f_1 and f_2 are functions. Since equations E_1 and E_2 are indeed functions whose respective inputs are u and v, it is not true that E_3 will uniquely determinate v when the values of b and u are given. This is only valid if guard b is true. Thus the directed graph (35) does not help finding the index.

What was wrong? The point is that, for non-smooth equations, it is not true that any variable involved in it can be assigned the status of output, even structurally. How can we fix this? The solution is obvious for this case: replace (35) by the following *guarded* consistent causality, in which predicate $b = [x{>}0]$ acts as a guard:

$$\begin{array}{l} x \rightarrow E_0 \rightarrow b \\ b \rightarrow E_3 \\ \textbf{if } b \textbf{ then } u \rightarrow E_3 \rightarrow v \textbf{ else } v \rightarrow E_3 \rightarrow u \end{array} \tag{36}$$

The actual consistent causality depends on the value of guard b. Observe that, in this guarded consistent causality, b is evaluated prior to being used as a guard. So, our first observation is that causality analyses must be guard dependent.

2) Unilateral constraint: A second interesting example is that of a unilateral constraint:

$$\begin{array}{lll} E_1: & x^\bullet = f(x,u) \\ E_2: & 0 \leq g(x) \end{array} \tag{37}$$

No consistent causality can be found for (37). The first difference array associated with (37) consists in adding the latent constraint $E_2^\bullet$, obtained by shifting E_2 forward:

$$\begin{array}{lll} E_1: & x^\bullet = f(x,u) \\ E_2: & 0 \leq g(x) \\ E_2^\bullet: & 0 \leq g(x^\bullet) \end{array} \tag{38}$$

Substituting $x^\bullet$ by $f(x, u)$ in $E_2^\bullet$, which we rename E_3, yields by setting $h(x, u) =_{\text{def}} g(f(x, u))$:

$$\begin{array}{lll} E_1: & x^\bullet = f(x,u) \\ E_2: & 0 \leq g(x) \\ E_3: & 0 \leq h(x,u) \end{array} \tag{39}$$

The consistent causality associated with (39) is thus the following, where equations have been reordered to match causality:

$$x \text{ consistent for } E_2 \tag{40}$$

$$x \rightarrow E_2^\bullet \rightarrow u \tag{41}$$

$$(x,u) \rightarrow E_1 \rightarrow x^\bullet \tag{42}$$

where "x consistent for E_2" means that $0 \leq g(x)$ and $(x, u){\rightarrow}E_1$ means $x{\rightarrow}E_1, u{\rightarrow}E_1$; (41) abstracts the unilateral constraint $E_2^\bullet$ acting on u and requires a solver. A valid execution scheme at each time step for (39) would be the following one, formalized as a constructive semantics:

Constructive Semantics 2:

1) (40): assume x consistent for E_2;
2) (41): use a suitable static constraint solver proposing a u such that $h(x, u) \geq 0$;
3) (42): having (x, u) compute $x^\bullet$, which is guaranteed consistent for the next time step. □

3) The need for atomic sets of equations: Here we continue the discussion of the example of unilateral constraint in Section III-E2 by developing a special but important point.

So far unilateral constraints are not part of our syntax of guarded equations. We could consider including them. There are, however, many more candidate primitives to be included—for instance the so-called complementarity constraints we discuss in the sequel. To avoid expanding our syntax with too

many primitives, can we instead expand unilateral constraints in terms of guarded equations? The answer is: yes!

$$g(x)\geq 0 \quad \text{expands as} \begin{cases} E_{21}: & b = [g(x)\leq 0] \\ E_{22}: & \textbf{if } b \textbf{ then } \underbrace{g(x)=0}_{F_{22}} \end{cases} \tag{43}$$

Unfortunately, the causality constraints induced by the set $\{E_{21}, E_{22}\}$ of equations has a circuit:

$$x \to E_{21} \to b \to E_{22}\,,\ \textbf{if } b \textbf{ then } F_{22} \to x \tag{44}$$

This should not come as a surprise since we know that (43) is a constraint. This leads to considering the set of equations $\{E_{21}, E_{22}\}$ as *atomic*, meaning that it must be evaluated at once. To reflect this on the causality analysis, simply perform the following transformations on (44):

1) Reinforce (44) to $x \to E_{21} \to b \to E_{22} \to x$;
2) Turn the so obtained circuit into the following consistent causality involving the atom $\{E_{21}, E_{22}\}(x, b)$:

$$\{E_{21}, E_{22}\} \to (x, b) \tag{45}$$

We now apply this technique to system (39), which expands as follows:

$$\begin{array}{rl} E_1: & x^\bullet = f(x,u) \\ E_{21}: & b = [0 \geq g(x)] \\ E_{22}: & \textbf{if } b \textbf{ then } 0 = g(x) \\ E_{31}: & c = [0 \geq h(x,u)] \\ E_{32}: & \textbf{if } c \textbf{ then } \underbrace{0 = h(x,u)}_{F_{32}(x,u)} \end{array} \tag{46}$$

An attempt for the consistent causality of (46) is:

$$(x,u) \to E_1 \to x^\bullet \tag{47}$$

$$x \text{ consistent for } \{E_{21}, E_{22}\}, E_{21} \to b \tag{48}$$

$$(x,u) \to E_{31} \to c \to E_{32},\ \textbf{if } c \textbf{ then } (x,u) \to F_{32} \to u \tag{49}$$

Causality constraints collected in (49) were derived by abstracting each individual equations as its causality constraints. Subset of equations $\{E_{31}, E_{32}\}$ exhibits a causality circuit, so we consider it as atomic. Applying to (49) the same transformation as we did from (44) to (45), yields:

$$x \to \{E_{31}, E_{32}\} \to (u, c) \tag{50}$$

thus resulting in the following consistent causality for (46):

$$\begin{array}{l} (x,u) \to E_1 \to x^\bullet \\ (x,b) \text{ consistent for } \{E_{21}, E_{22}\} \\ x \to \{E_{31}, E_{32}\} \to (u,c) \end{array}$$

which coincides with (40)–(42) if we abstract b and c away. In reinforcing (49) to (50), we decided to have x as the source and u as a target. We could have equally well done the opposite:

$$u \to \{E_{31}, E_{32}\} \to (x, c) \tag{51}$$

In other words, the atom $\{E_{31}, E_{32}\}$ is assigned two possible causality constraints, namely

$$\begin{array}{ll} & x \to \{E_{31}, E_{32}\} \to (u,c) \\ \text{or} & u \to \{E_{31}, E_{32}\} \to (x,c). \end{array} \tag{52}$$

Which one is ultimately kept depends on the other causality constraints, when attempting to avoid circuits. To summarize this discussion, when dealing with a minimal circuit in the set of causality constraints, the following must be performed:

1) propose for it a set of candidate circuitfree causality constraints;
2) taking the causality constraints induced by the whole set of equations into account, select a compatible candidate (causing no global circuit), if any.

In this work, we are not proposing a systematic procedure for implementing Task 1). We can see 1) as manual task to be performed for a given library of primitive constraints (e.g., the unilateral constraint and the complementarity constraint discussed below). Grouping to an atom the equations involved in a circuit and applying the above procedure allows invoking "local solvers" as part of the execution of dAE.

4) Complementarity condition: The following is a simple case of non-smooth dAE system, namely a dynamical system subject to a so-called *complementarity condition*, usually written as $0 \leq U(x) \perp V(y) \geq 0$:

$$\begin{array}{r} U(x)\geq 0 \text{ and } V(y)\geq 0 \text{ and } U(x)V(y) = 0 \\ F(x,y) = 0 \end{array} \tag{53}$$

where x and y have dimensions possibly >1, U and V are $\mathbb{R}$-valued functions, and the constraint $F(x,y)=0$ is assumed to make the overall system (53) nonsingular. Such systems are frequently encountered in circuits (e.g., perfect diodes) and mechanics (e.g., dry friction, or contact). The complementary condition is a multi-mode constraint. We reformulate example (53) using guarded equations as follows:

$$\begin{array}{rl} E_1: & 0 = F(x,y) \\ E_{21}: & b_U = [U(x) > 0] \\ E_{22}: & b_V = [V(y) > 0] \\ E_{23}: & \textbf{if } b_U \textbf{ then } 0 = V(y) \textbf{ else } 0 = U(x) \\ E_{24}: & \textbf{if } b_V \textbf{ then } 0 = U(x) \textbf{ else } 0 = V(y) \end{array} \tag{54}$$

Regard $E_2 =_{\text{def}} \bigwedge_{1\leq i\leq 4} E_{2i}$ as an atomic system. Its consistent causality is derived similarly as for system (46), where E_U and E_V denote the equations $0=U(x)$ and $0=V(y)$, respectively:

$$\begin{array}{l} x \to E_{21} \to b_U \\ y \to E_{22} \to b_V \\ b_U \to E_{23}, \textbf{if } b_U \textbf{ then } E_V \to y \textbf{ else } E_U \to x \\ b_V \to E_{24}, \textbf{if } b_V \textbf{ then } E_U \to x \textbf{ else } E_V \to y \end{array} \tag{55}$$

In (55) we are missing the important side information that $b_U \wedge b_V = \text{F}$. We add it as an assertion:

$$\begin{array}{l} \textbf{assert } b_U \wedge b_V = \text{F} \\ x \to E_{21} \to b_U \\ y \to E_{22} \to b_V \\ b_U \to E_{23}, \textbf{if } b_U \textbf{ then } E_V \to y \textbf{ else } E_U \to x \\ b_V \to E_{24}, \textbf{if } b_V \textbf{ then } E_U \to x \textbf{ else } E_V \to y \end{array} \tag{56}$$

We are now ready to include E_1 in our causality analysis. The entire system (54) is now considered atomic, with the

following set of causality constraints:

$$\begin{array}{l}\textbf{assert } b_U \wedge b_V = \text{F}\\ x{\rightarrow}E_{21}{\rightarrow}b_U\\ y{\rightarrow}E_{22}{\rightarrow}b_V\\ b_U{\rightarrow}E_{23}, \textbf{if } b_U \textbf{ then } E_V{\rightarrow}y{\rightarrow}E_1{\rightarrow}x\\ \qquad\qquad\qquad\quad \textbf{else } E_U{\rightarrow}x{\rightarrow}E_1{\rightarrow}y\\ b_V{\rightarrow}E_{24}, \textbf{if } b_V \textbf{ then } E_U{\rightarrow}x{\rightarrow}E_1{\rightarrow}y\\ \qquad\qquad\qquad\quad \textbf{else } E_V{\rightarrow}y{\rightarrow}E_1{\rightarrow}x\end{array} \tag{57}$$

Summary of this informal discussion:

- We will systematically expand unilateral constraints into guarded equations;
- While doing so, we may need to "glue together" some equations and see them as an *atomic system*;
- Atomic systems are given at once a consistent causality. This consistent causality may involve assertions relating guards, which may somtimes be abstracted as a nondeterministic choice between alternatives.

In the next section we formalize the above approach.

IV. Index of non-smooth DAE: theory

In a first stage, we ignore guards. So the causality analyses we develop will be guardless. This will raise difficulties as we shall see. Thus, in a second stage, we account for guards.

A. Guardless causality analysis

We assume an underlying set of variables, together with the generic notation for its elements:

$$x \in \mathcal{X}$$

For $x \in \mathcal{X}$, we denote by $D(x)$ the domain of x and we set $D(\mathcal{X}) = \prod_{x\in\mathcal{X}} D(x)$. Then, for k any positive integer, we consider a fresh copy

$$\mathcal{X}^{\bullet k}$$

of $\mathcal{X}$ and we write $\mathcal{X}^\bullet$ instead of $\mathcal{X}^{\bullet 1}$; the elements of $\mathcal{X}^{\bullet k}$ are $x^{\bullet k}$. In particular, $D(x^\bullet) = D(x)$ and $D(\mathcal{X}^\bullet) = D(\mathcal{X})$. Next, we assume an underlying set

$$\mathbb{E}$$

of primitive *constraints* $E \subseteq D(\mathcal{X})\times D(\mathcal{X}^\bullet)$, which we call *equations* in the sequel. Some subsets of this set of equations can be specified as being *atomic*. Atomic subsets are assumed to be pairwise disjoint. Atoms consisting of a singleton are identified with the single primitive equation they contain. Thus, for our theoretical development, $\mathbb{E}$ will be an underlying set of atoms generically denoted by the symbol E. For convenience, the term "equation" will often be used instead of "atom" in the sequel.

The dAE systems we consider are finite conjunctions of equations $C = \bigwedge_{i\in I} E_i$. For convenience, we shall often identify system C with its set $\{E_i | i{\in}I\}$ of constitutive equations. Write

$$x \overline{\in} E$$

to indicate that x occurs in E and denote by

$$\overline{\in}_E$$

the set of all variables x such that $x \overline{\in} E$. If E is an equation of the form $F = 0$ where F is smooth, $x \overline{\in} E$ is equivalent to saying that the partial derivative $\partial F/\partial x$ is structurally nonzero. For $E \in \mathbb{E}$ and $x \overline{\in} E$, we consider the predicate

$$x \notin_{\text{out}} E \tag{58}$$

to specify that the value of x must be known prior to solving equation E (intuitively, x cannot be seen as an output of equation E).

Comment 1: Predicate (58) is not a semantic property of equation E; it is rather an additional specification that must be provided together with the semantics of E. For example, for $E : x{+}y{=}0$, neither x nor y satisfies (58). For $E : b{=}[x > 0]$, we specify $x \notin_{\text{out}} E$. For the guarded equation

$$E : \textbf{if } b \textbf{ then } F_1(u, v){=}0 \textbf{ else } F_2(u, v){=}0$$

we specify $b \notin_{\text{out}} E$. For the atom $E =_{\text{def}} \{E_{31}, E_{32}\}$ of system (46), we specify $c \notin_{\text{out}} E$, which is equivalent to the set (52) of two candidate causalities. □

For C a dAE system, define

$$x \overline{\in} C \text{ if and only if } x \overline{\in} E \text{ holds for some } E{\in}C$$

and denote by $\overline{\in}_C$ the set of variables $z \in \mathcal{X}{\uplus}\mathcal{X}^\bullet$ such that $z \overline{\in} C$. Define, for k a non-negative integer:

$$\begin{array}{lll}\mathcal{X}_C & =_{\text{def}} & \overline{\in}_C \cap \mathcal{X}\\ \mathcal{V}_C & =_{\text{def}} & \overline{\in}_C \cap \mathcal{X}^\bullet\\ \mathcal{W}_C & =_{\text{def}} & \{x^{\bullet l} \,|\, 1{\leq}l{\leq}k, x^\bullet \notin \mathcal{V}_C, x \in \mathcal{X}_C\} \cup\\ & & \{x^{\bullet l} \,|\, 2{\leq}l{\leq}k, x \in \mathcal{V}_C\}\\ \mathbb{X}_C & =_{\text{def}} & \mathcal{X}_C \uplus \mathcal{V}_C \uplus \mathcal{W}_C\end{array} \tag{59}$$

Variables belonging to $\mathcal{V}_C$ are the state variables of C and variables belonging to $\mathcal{X}_C$ are its other variables. Variables belonging to $\mathcal{W}_C$ are the additional variables resulting from shifting forward the former ones and integer k is the given bound for those shifts. The following definition is a variation of Definition 4 to account for specified causality constraints:

Definition 7: Let C' be a dAE system and $C = \bigwedge_{i\in I} E_i$ be its k-th difference array following (27). We associate to array C its Pantelides graph *$\mathcal{P}_C$, which is a bipartite graph defined by the following rules, where we identify C with its set of constitutive equations $\{E_i \mid i{\in}I\}$:*

- *Its set of vertices is $\mathbb{X}_C \uplus C$;*
- *For every $z \in \mathbb{X}_C$, using notation (22):*

$$\frac{E \in C \text{ and } z \overline{\in} E}{z{-}E \in \mathcal{P}_C}$$

Call consistent causality *of C any directed bipartite graph $\vec{\mathcal{P}}_C$ such that*

1) *Its set of vertices is $\mathcal{Z} \uplus C$, for some $\mathcal{Z} \subseteq \mathbb{X}_C$ such that $\mathcal{X}_C \uplus \mathcal{V}_C \subseteq \mathcal{Z}$;*

2) *$\vec{\mathcal{P}}_C$ covers $\mathcal{P}_C$: for every $z \in \mathcal{Z}$,*

$$\frac{z\text{—}E \in \mathcal{P}_C}{\text{either} \quad z\rightarrow E \in \vec{\mathcal{P}}_C \quad \text{or} \quad E\rightarrow z \in \vec{\mathcal{P}}_C}$$

3) *Causality specifications are respected:*

$$\frac{z\text{—}E \in \mathcal{P}_C \text{ and } z \notin_{\text{out}} E}{z\rightarrow E \in \vec{\mathcal{P}}_C}$$

4) *$\vec{\mathcal{P}}_C$ has the* single assignment *property, meaning that, for every z, there exists at most one equation E such that $E\rightarrow z \in \vec{\mathcal{P}}_C$;*
5) *$\vec{\mathcal{P}}_C$ is circuitfree; its transitive closure $\preceq_C$ is thus a partial order on $\mathcal{Z} \uplus C$, called the* causality order *induced by $\vec{\mathcal{P}}_C$;*
6) *Denoting by $\min(\preceq_C)$ the set of the minimal elements of $\preceq_C$, we have $\min(\preceq_C) \subseteq \mathcal{X}_C$;*
7) *For every pair $(z, w) \in (\mathcal{X}_C \uplus \mathcal{V}_C) \times (\mathcal{W}_C \cap \mathcal{Z})$, it never holds that $w \preceq_C z$.*

The index *of dAE system C' is re-defined as the minimal integer k such that its k-th array C possesses a consistent causality.*

The novelty with respect to Definition 4 lies in condition 3), which accounts for the fact that, in searching for a consistent causality, some constraints are subject to causality specifications, see Comment 1. Definition 4 is justified by the following weakening of Theorem 1, which follows from a straightforward pivoting argument:

Theorem 2: Consider the following condition for the considered consistent causality:

$$\begin{array}{l} \textit{for every branch } E\rightarrow z \in \vec{\mathcal{P}}_C,\ z \textit{ is} \\ \textit{uniquely determined as an output of } E, \\ \textit{for any consistent choice of values} \\ \textit{for the other variables involved in } E. \end{array} \tag{60}$$

Assuming (60), if C possesses a consistent causality, then property (31) holds, and thus C possesses a constructive semantics.

If the system is smooth, condition (60) holds structurally. The example (34)–(36) in section III-E illustrated how condition (60) may be violated by non-smooth systems. See the discussion sitting between (35) and (36), where the use of guards was motivated. In the next section, we formalize this by refining our development with the consideration of guards. In passing, we describe the constructive semantics, which yields the execution scheme.

B. Guarded causality analysis

We assume a subset $\mathcal{B} \subset \mathcal{X}$ of *guards* and we augment $\mathbb{E}$ with two new kinds of equations:

Assertions over guards:

$$\textbf{assert}\ P(b_1, \ldots, b_k) \quad \text{where} \quad b_1, \ldots, b_k \in \mathcal{B} \tag{61}$$

where $P(b_1, \ldots, b_k)$ is a boolean expression over the listed guards; this assertion states that predicate $P(b_1, \ldots, b_k)$ holds true in the dAE system where this assertion occurs.

Guarded equations:

$$\textbf{if}\ b\ \textbf{then}\ E \tag{62}$$

where $(b, E) \in \mathcal{B} \times \mathbb{E}$; this guarded equation specifies that equation E is in force whenever guard b is true.

Convention and associated notation: We use the generic notation

$$\mathbf{E} \tag{63}$$

to denote a *possibly guarded equation,* i.e., a guarded equation (61) or simply an element of $\mathbb{E}$. For $\mathbf{E}$ a guarded equation of the form "**if** b **then** E", we denote by

$$b(\mathbf{E}) =_{\text{def}} b$$

its guard. For a non guarded equation E and considering Axiom 1 below, we set by convention $b(E) = \text{T}$.

In the sequel, dAE systems are pairs consisting of a conjunction of possibly guarded equations, and a conjunction of assertions over the involved guards:

$$C = \left(\bigwedge_{i \in I} \mathbf{E}_i\ ,\ \bigwedge_{j \in J} \textbf{assert}\ P_j \right) \tag{64}$$

To properly capture the intuition behind (62), we assume the following about this operation:

Axiom 1: The following holds:

$$\forall E \in \mathbb{E} \quad \Longrightarrow \quad [\,\textbf{if}\ \text{T}\ \textbf{then}\ E\,] \equiv E \tag{65}$$
$$\forall E \in \mathbb{E} \quad \Longrightarrow \quad [\,\textbf{if}\ \text{F}\ \textbf{then}\ E\,] \equiv \epsilon \tag{66}$$
$$\forall \mathbf{E} \in \mathbb{E} \quad \Longrightarrow \quad b(\mathbf{E}) \notin_{\text{out}} \mathbf{E} \tag{67}$$

(65) states that, if b is true, equation "**if** b **then** E" reduces to E. (66) states that, if the guard b is false, then "**if** b **then** E" collapses to the trivial equation ϵ setting no constraint at all, i.e., having no variable involved in it: $\overline{\in}_\epsilon = \emptyset$. Focus finally on condition (67) by making the form of $\mathbf{E}$ explicit: $\mathbf{E}$: "**if** b **then** E". Then, condition (67) expresses that guard b must be evaluated prior to the equation it guards, see (58) for the definition of $\notin_{\text{out}}$. Observe that this is a nontrivial restriction, since it prevents **if** b **then** E from being a fixpoint in both b and the variables of E, jointly.

We now formalize the notion of guarded Pantelides graph, which provide the graph abstraction of guarded equations. For $(b, x, E) \in \mathcal{B} \times \mathcal{X} \times \mathbb{E}$, we consider the following *guarded bipartite branch* and guarded *directed* bipartite branch:

$$\textbf{if}\ b\ \textbf{then}\ x\text{—}E \quad \text{and} \quad \left\{ \begin{array}{ll} & \textbf{if}\ b\ \textbf{then}\ x\rightarrow E \\ \text{or} & \textbf{if}\ b\ \textbf{then}\ E\rightarrow x \end{array} \right. \tag{68}$$

(We make no difference between non directed branches $x\text{—}E$ and $E\text{—}x$.) We consider the counterpart of Axiom 1:

Axiom 2: The following holds, for any (possibly directed) bipartite branch π:

$$[\,\textbf{if}\ \text{T}\ \textbf{then}\ \pi\,] \quad \equiv \quad \pi \tag{69}$$
$$[\,\textbf{if}\ \text{F}\ \textbf{then}\ \pi\,] \quad \equiv \quad \epsilon \tag{70}$$

Due to Axiom 2, we feel free to identify the non guarded branch π with the guarded branch "**if** T **then** π". For $\pi =$

if b **then** $x—E$ (respectively "**if** b **then** $x\rightarrow E$") a branch, define its *base*:

$$\begin{array}{rrcl} & \lfloor\pi\rfloor & =_{\rm def} & x—E \\ \text{respectively} & \lfloor\pi\rfloor & =_{\rm def} & x\rightarrow E \end{array}$$

and, for the directed branch, its *reversal*:

$$\overleftarrow{\pi} \quad =_{\rm def} \quad \textbf{if } b \textbf{ then } E\rightarrow x$$

Due to Axiom 2, we feel free to identify the non guarded branch π with the guarded branch "**if** T **then** π". Accordingly, for a non guarded branch π, its base $\lfloor\pi\rfloor$ coincides with π itself.

A *guarded (directed) bipartite graph* $\mathcal{P}$ is a pair consisting of a finite set of guarded (directed) bipartite branches, and a finite conjunction of assertions over its involved guards:

$$\mathcal{P} = \left(\{\pi_i \mid i\in I\}\,,\, \bigwedge_{j\in J} \textbf{assert } P_j\right) \tag{71}$$

In the sequel, we simply write $\pi \in \mathcal{P}$ to indicate that π is a branch of the graph part of $\mathcal{P}$.

If $\vec{\mathcal{P}}$ is directed, say that $\pi_1, \ldots, \pi_k$ is a *path* (*circuit*) of $\vec{\mathcal{P}}$ if $\lfloor\pi_1\rfloor, \ldots, \lfloor\pi_k\rfloor$ is a path (circuit). For $\pi_1, \ldots, \pi_k$ a path in $\vec{\mathcal{P}}$ such that $\pi_i = \textbf{if } b_i \textbf{ then } \lfloor\pi_i\rfloor$, we define its *guard* $b(\pi_1, \ldots, \pi_k)$ by

$$b(\pi_1, \ldots, \pi_k) \quad =_{\rm def} \quad \bigwedge_{1\leq i\leq k} b_i \tag{72}$$

and say that $\vec{\mathcal{P}}$ is *guarded circuitfree* if all its circuits have a false guard—taking assertions of $\vec{\mathcal{P}}$ into account. If $\vec{\mathcal{P}}$ is circuitfree, then the transitive closure $\preceq_{\vec{\mathcal{P}}}$ of the graph part of $\vec{\mathcal{P}}$ is a partial order, for any configuration of the guards.

Say that $\vec{\mathcal{P}}$ has the *guarded single assignment* property if the assertions of $\vec{\mathcal{P}}$ imply that, for every z, $\bigwedge_{\ell\in L} b_\ell = \mathrm{F}$ holds, where $\{b_\ell \mid \ell\in L\}$ is the set of guards of all the branches "**if** b_ℓ **then** $E_\ell\rightarrow z$" of $\vec{\mathcal{P}}$ ending at z. Using these prerequisites, Definition 7 is reformulated:

Definition 8: Let C' be a dAE system and

$$C \quad = \quad \left(\bigwedge_{i\in I} \mathbf{E}_i\,,\, \bigwedge_{j\in J} \textbf{assert } P_j\right)$$

be its k-th difference array following (27), together with its set of assertions. We associate to array C its guarded Pantelides graph *$\mathcal{P}_C$, which is a guarded bipartite graph keeping the assertions of C and otherwise obtained by applying the following rules, where we identify C with its set of constitutive equations $\{\mathbf{E}_i \mid i\in I\}$:*

$$\begin{array}{c} \dfrac{\mathbf{E}\in C}{[\,b(\mathbf{E})—\mathbf{E}\,]\in\mathcal{P}_C} \\[2ex] \dfrac{[\,\textbf{if } b \textbf{ then } E\,]\in C \quad \text{and} \quad z\,\overline{\in}\,E}{[\,\textbf{if } b \textbf{ then } z—E\,]\in\mathcal{P}_C} \end{array} \tag{73}$$

Call guarded consistent causality *of C any guarded directed bipartite graph $\vec{\mathcal{P}}_C$ keeping the assertions of C and otherwise such that, with reference to notations (59):*

1) *Its set of vertices is $\mathcal{Z}\uplus C$, for some $\mathcal{Z}\subseteq\mathbb{X}_C$ such that $\mathcal{X}_C\uplus\mathcal{V}_C\subseteq\mathcal{Z}$;*
2) *$\vec{\mathcal{P}}_C$ covers $\mathcal{P}_C$: for every π,*
$$\frac{\pi\in\mathcal{P}_C}{\text{either}\quad \pi\in\vec{\mathcal{P}}_C \quad\text{or}\quad \overleftarrow{\pi}\in\vec{\mathcal{P}}_C}$$
3) *Causality specifications are respected:*
$$\frac{[\textbf{if } b \textbf{ then } z—E]\in\mathcal{P}_C \quad\text{and}\quad z\notin_{\rm out} E}{[\textbf{if } b \textbf{ then } z\rightarrow E]\in\vec{\mathcal{P}}_C}$$
4) *Causality distributes over guards:*
$$\frac{[\,z\rightarrow[\,\textbf{if } b \textbf{ then } E\,]\,]\in\vec{\mathcal{P}}_C}{[\,\textbf{if } b \textbf{ then } z\rightarrow E\,]\in\vec{\mathcal{P}}_C}$$
5) *$\vec{\mathcal{P}}_C$ has the guarded single assignment property;*
6) *$\vec{\mathcal{P}}_C$ is guarded circuitfree, hence we can consider the transitive closure $\preceq_C$ of the graph part of $\vec{\mathcal{P}}_C$, which we call the* guarded causality order *induced by $\vec{\mathcal{P}}_C$;*
7) *Denoting by $\min(\preceq_C)$ the set of the vertices that are minimal elements of $\preceq_C$, then $\min(\preceq_C)\subseteq\mathcal{X}_C\uplus C'$ holds;*
8) *For every pair $(z,w)\in(\mathcal{X}_C\uplus\mathcal{V}_C)\times(\mathcal{W}_C\cap\mathcal{Z})$, it never holds that $w\preceq_C z$.*

The index *of dAE system C' is re-defined as the minimal integer k such that the k-th difference array C possesses a consistent causality.*

Observe that, due to condition (67) of Axiom 1, entailment 3), applied with b being the constant *true* and E replaced by $\mathbf{E} =$ "**if** b **then** E", yields:

$$\frac{\mathbf{E}\in C}{[\,b(\mathbf{E})\rightarrow\mathbf{E}\,]\in\vec{\mathcal{P}}_C} \tag{74}$$

(Compare with the first entailment rule in (73).)

C. Constructive semantics

The following result justifies the consideration of Definition 8:

Theorem 3: Assuming (60), if C possesses a consistent causality according to Definition 8, then property (31) holds.

Said differently, Theorem 2 still holds with Definition 8 substituted for Definition 7. The important point about refining Definition 7 into Definition 8 is that condition (60), which ensures the correctness of the graphical abstraction, is much easier satisfying, see the paragraph sitting just before Section IV-B.

Proof: By condition 7) and entailment (74), vertices belonging to $\min(\preceq_C)$ can only be of two categories:

- either non guarded equations having no inputs and only outputs; such equations belong to $\min(\preceq_C)$ whatever configuration of the guards is;
- or variables belonging to $\mathcal{X}_C$ that are output of no equation; such variables belong to $\min(\preceq_C)$ whatever configuration of the guards is.

In particular, the set $\min(\preceq_C)$ is independent of the configuration of the guards and, thus, condition 7) is indeed meaningful.

The evaluation of all variables of C at a considered instant proceeds according to the following constructive semantics:

Constructive Semantics 3: $\vec{\mathcal{P}}$ is a running subgraph of $\vec{\mathcal{P}}_C$ and $\preceq$ a running subrelation of $\preceq_C$:

1) Initialization: $\vec{\mathcal{P}} := \vec{\mathcal{P}}_C$ and $\preceq := \preceq_C$;
2) While $\vec{\mathcal{P}}$ has non-empty set of vertices, do:
   a) evaluate variables or equations belonging to $\min(\preceq)$; the variables are evaluated by reading their values; thanks to Condition 5) of Definition 8 and condition (60), evaluating the equations fixes the values of their immediate sucessors in $\preceq$;
   b) for those evalued variables that are guards, replace them by their value T or F and then apply simplifying rules (65) and (66) of Axiom 1; Condition 4 of Definition 8 ensures that no causality is lost by removing the trivial guards T;
   c) erase, from $\vec{\mathcal{P}}$, the vertices belonging to $\min(\preceq)$ and adjacent edges and redefine $\preceq$ accordingly. □

The key observation is that, at each round of Algorithm 3, $\min(\preceq)$ is independent of the configuration of the guards that remain to be evaluated. By Condition 2) of Definition 8, every equation and variable of array C gets evaluated. Finally, due to Condition 8) of Definition 8, unnecessary variables w are eliminated by discarding the equation defining them. This finishes the proof of Theorem 3. ■

Constructive Semantics 3 can be seen as an interpreter of system C.

Comment 2: Entailment rule (74) forbids the consideration of systems in which guards are themselves solutions of fixpoint equations, such as in the following example:

$$\begin{cases} b = [\, x > 0\,] \\ \textbf{if } b \textbf{ then } [\, 0 = f(x)\,] \end{cases}$$

which means informally: **if** b **then** $f(x)$=0 **where** b=[x>0]. Such fixpoint equations cannot be solved by graph based abstractions. Numerical methods must be used—a natural candidate being relaxation, by which a sequence $b_0, x_1, b_1, x_2, \ldots$ of values is computed by using the two equations in a loop until (hopefully) fixpoint occurs. Graph based algorithms *à la* Pantelides do require condition (67) of Axiom 1.

Comment 3 (atoms): Recall that some of the E's involved in the considered dAE system are atomic systems of equations representing constraints for which a solver is needed. The principle is that each such constraint E comes with its guarded Pantelides graph $\mathcal{P}_E$, possibly enhanced with a set of assertions on the guards of E and a set of specifications of the form $x \notin_{\text{out}} E$, see (58). By using the context of E, if a consistent causality can be found, a directed graph $\vec{\mathcal{P}}_E$ results, which in turn specifies the variables for which the solver must return a solution of E. See Section III-E3 for a detailed example.

Comment 4 (nested guards): Capturing multi-mode systems with arbitrarily nested modes would require that $\mathbb{E}$ is further extended to become closed under the mapping $(b, E) \mapsto [\textbf{if } b \textbf{ then } E]$, meaning that nested guards should be considered—something forbidden by our current extension of $\mathbb{E}$, see (62). Extending Definition 8 and Theorem 3 to cover this is feasible but more involved. This is left for further work.

Comment 5 (multiple clocked dAE systems): So the dAE as defined in Section III-A are "single-clocked" in that all variables possess a value at every instant. It is well known from synchronous languages that multiple-clocked discrete time systems can be made multiple-clocked by adding, to every data type, a distinguished symbol "$\perp$" denoting the absence of the considered variable at the considered instant [5]. In a given run, the subset of instants at which a given variables is present, i.e., $\neq\perp$, is called its *clock*. Clocks can be seen as a type system and these types can either be verified or synthesized [5]. Of course, the same techniques apply to dAE systems. We did not consider this issue here since it would bring yet another layer of technicality.

D. Back to dAE examples

In this section we revisit the non-smooth dAE examples of Section III-E and we add the most basic example, namely the change of mode upon detection of a zero-crossing.

1) Zero-crossing: The following example of continuous-time ODE with event-based reset is the simplest case of hybrid system, i.e., continuous time system with mode change:

$$\textbf{if } b \textbf{ then } x = h(x^-) \textbf{ else } \dot{x} = f(x), \tag{75}$$

where boolean signal b selects the events of *zero-crossings* of some smooth function $g(x)$, i.e., at any instant when g crosses zero from below. The default dynamics is the ODE $\dot{x} = f(x)$ and $x = h(x^-)$ yields the reset value for x at the instants of zero-crossing, where x^- denotes the left-limit of trajectory x, i.e., the value x had just before the reset.

We consider here example (75), albeit in its dAE form. Three variations can be considered, depending on how shifts get substituted for the left-limit and derivative operators:

$$\begin{cases} \textbf{if } b \textbf{ then } x = h({}^\bullet x) \textbf{ else } x^\bullet = f(x) \\ b = Q(x, x^\bullet) \end{cases} \tag{76}$$

$$\begin{cases} \textbf{if } b \textbf{ then } x = h({}^\bullet x) \textbf{ else } x^\bullet = f(x) \\ b = Q({}^\bullet x, x) \end{cases} \tag{77}$$

$$\begin{cases} \textbf{if } b \textbf{ then } x^\bullet = h(x) \textbf{ else } x^\bullet = f(x) \\ b = Q({}^\bullet x, x) \end{cases} \tag{78}$$

where Q is the predicate specifying the zero-crossings of g:

$$Q(z, x) \quad =_{\text{def}} \quad [g(z) \leq 0] \wedge [g(x) > 0] \tag{79}$$

We successively analyze the above three variants.

Analyzing (76)*:* Since (76) involves two successive instants, we first put it in state space form:

$$\begin{cases} z^\bullet = x \\ \textbf{if } b \textbf{ then } x = h(z) \textbf{ else } x^\bullet = f(x) \\ b = Q(x, x^\bullet) \end{cases} \tag{80}$$

which possesses no consistent causality and is, therefore, not of index 0. Denoting by $E(b, z, x, x^\bullet)$ the guarded equation

of (80), the next difference array for consideration is the following, where the "− − − −" indicate the separation between the original and shifted dynamics:

$$\left[\begin{array}{l} 0 = z^\bullet - x \\ E(b, z, x, x^\bullet) \\ 0 = b - Q(x, x^\bullet) \\ - - - - - - - - - - \\ 0 = z^{\bullet 2} - x^\bullet \\ E(b^\bullet, z^\bullet, x^\bullet, x^{\bullet 2}) \\ 0 = b^\bullet - Q(x^\bullet, x^{\bullet 2}) \end{array}\right]$$

According to the definition of the 2-nd array, variables $b^\bullet, x^{\bullet 2}, z^{\bullet 2}$ have to be eliminated. Unfortunately, considering this array and even further increasing it does not help finding a consistent causality for the pair $(b, x^\bullet)$. The index is thus found to be infinite. Let us try to fix this problem by specifying the predicate evaluation to be directed, see (58) and Comment 1 for the definition of $\not\in_{\text{out}}$:

$$\left\{\begin{array}{lll} & & 0 = z^\bullet - x \\ & & E(b, z, x, x^\bullet) \\ (x, x^\bullet) \not\in_{\text{out}} & \text{in} & b = Q(x, x^\bullet) \end{array}\right.$$

As a consequence, we cannot select $x^\bullet$ as an output of the first equation, since this would yield a causality circuit. The conclusion is that (77) is a bad model for a dAE version of zero-crossing.

Analyzing (77)*:* Repeating the same for (77) and using previous notation $E(b, z, x, x^\bullet)$ yields:

$$\left\{\begin{array}{lll} E_1: & & 0 = z^\bullet - x \\ E_2: & & E(b, z, x, x^\bullet) \\ E_3: & (z, x) \not\in_{\text{out}} \text{ in } & b = Q(z, x) \end{array}\right. \tag{81}$$

At this point we make an attempt to use guardless causality analysis, based on Definition 7. According to this definition, the following is a correct consistent causality for (81):

$$\left\{\begin{array}{r} x \to E_1 \to z^\bullet \\ (b, z, x) \to E_2 \to x^\bullet \\ (z, x) \to E_3 \to b \end{array}\right. \tag{82}$$

Consistent causality (82) is optimistic and erroneous, however, in that it does not satisfy the important condition (60). While it yields a seemingly correct order, it makes the assumption that $x^\bullet$ can be determined as an output of equation E_2. This is, however, wrong at zero-crossing events, i.e., when $b = \text{T}$. Thus, to guarantee that Theorem 3 applies, Definition 8 must be invoked and, thus, the following guarded Pantelides graph must be used instead (note the non-directed branches in the third and fourth lines):

$$\left\{\begin{array}{l} x \to E_1 \to z^\bullet \\ b \to E_2 \\ \textbf{if } b \textbf{ then } b \to E_{2,\text{T}} - (z, x) \\ \quad \textbf{else } \; b \to E_{2,\text{F}} - (x, x^\bullet) \\ (z, x) \to E_3 \to b \end{array}\right. \tag{83}$$

where $E_{2,\text{T}}$ and $E_{2,\text{F}}$ denote the if-branch and else-branch of guarded equation E_2. Unfortunately, no consistent causality can be found for (83) since every attempt is not circuitfree by exhibiting the circuit $x \to b \to x$ having guard b, which is not false in general. Thus, (77) is not appropriate either.

Analyzing (78)*:* A cascade of two successive events occurs, namely: the observation of $b = \text{T}$ at some instant, followed by a reset of x at the next instant. Putting (78) in state space form yields:

$$\left\{\begin{array}{l} z^\bullet = x \\ \textbf{if } b \textbf{ then } x^\bullet = h(x) \textbf{ else } x^\bullet = f(x) \\ b = Q(z, x) \end{array}\right. \tag{84}$$

for which the following consistent causality is found, showing that the index is 0:

$$\left\{\begin{array}{l} x \to E_1 \to z^\bullet \\ b \to E_2 \\ \textbf{if } b \textbf{ then } (b, x) \to E_{2,\text{T}} \to x^\bullet \textbf{ else } (b, x) \to E_{2,\text{F}} \to x^\bullet \\ (z, x) \to E_3 \to b \end{array}\right.$$

We can simplify the guarded branch by merging its two alternatives, which finally yields:

$$\left\{\begin{array}{l} x \to E_1 \to z^\bullet \\ (b, x) \to E_2 \to x^\bullet \\ (z, x) \to E_3 \to b \end{array}\right.$$

This is an example of the manipulations we can perform on guarded branches to avoid enumerating modes while computing the guarded incidence matrix of Pantelides' graph.

2) Unilateral constraint: The unilateral constraint in discrete time was studied in Sections III-E2 and III-E3. The developments performed there, exactly follow the theory of Definition 8, Theorem 3, and Constructive Semantics 3.

3) Complementarity condition : Complementarity condition (53) of Section III-E4 is reformulated as the following dAE system with guards and assertions:

$$\begin{array}{ll} E_1: & 0 = F(x, y) \\ A_2: & \textbf{assert } b_U \wedge b_V = \text{F} \\ E_{21}: & b_U = [U(x) > 0] \\ E_{22}: & b_V = [V(x) > 0] \\ E_{23}: & \textbf{if } b_U \textbf{ then } 0 = V(y) \textbf{ else } 0 = U(x) \\ E_{24}: & \textbf{if } b_V \textbf{ then } \underbrace{0 = U(x)}_{F_{24}} \textbf{ else } \underbrace{0 = V(y)}_{F_{23}} \end{array} \tag{85}$$

With reference to (53), we have added the assertion that guards b_U and b_V are never true simultaneously. Clearly, this can be deduced from the equations defining the system. Still, we state it explicitly so it can be used when constructing the causality analysis.

No consistent causality can be found for it other than by viewing the entire system (85) but its first equation as an atom. Doing so is indeed consistent with (85) being itself, in the litterature of so-called non-smooth systems, seen as a primitive for which dedicated solvers are developed [1]. Depending on the particular form for the first equation, we may need to extend (85) into a larger array, by shifting forward the atom.

The following set of causality constraints is found for the atom $(A_2, E_2) =_{\text{def}} (A_2, \{E_{21}, E_{22}, E_{23}, E_{24}\})$:

$$\begin{array}{ll} A_2: & \textbf{assert } b_U \wedge b_V = \text{F} \\ E_2: & \left\{ \begin{array}{l} x \rightarrow E_2 \rightarrow (b_U, b_V) \\ b_U \rightarrow E_{23} \\ \textbf{if } b_U \textbf{ then } b_U \rightarrow F_{23} \rightarrow y \textbf{ else } b_U \rightarrow F_{24} \rightarrow x \\ b_V \rightarrow E_{24} \\ \textbf{if } b_V \textbf{ then } b_V \rightarrow F_{24} \rightarrow x \textbf{ else } b_V \rightarrow F_{23} \rightarrow y \end{array} \right. \end{array}$$

which boils down to

$$\begin{array}{ll} A_2: & \textbf{assert } b_U \wedge b_V = \text{F} \\ E_2: & \left\{ \begin{array}{l} x \rightarrow E_2 \rightarrow (b_U, b_V) \\ b_U \rightarrow E_{23}, \textbf{if } b_U \textbf{ else } b_U \rightarrow F_{24} \rightarrow x \\ b_V \rightarrow E_{24}, \textbf{if } b_V \textbf{ else } b_V \rightarrow F_{23} \rightarrow y \end{array} \right. \end{array} \tag{86}$$

Combining (86) with equation E_1 yields the following set of causality constraints for the overall atomic system (85):

$$\begin{array}{ll} A_2: & \textbf{assert } b_U \wedge b_V = \text{F} \\ \{E_1, E_2\}: & \left\{ \begin{array}{l} x \rightarrow E_2 \rightarrow (b_U, b_V) \\ b_U \rightarrow E_{23}, \textbf{if } b_U \textbf{ else } b_U \rightarrow F_{24} \rightarrow x \rightarrow E_1 \rightarrow y \\ b_V \rightarrow E_{24}, \textbf{if } b_V \textbf{ else } b_V \rightarrow F_{23} \rightarrow y \rightarrow E_1 \rightarrow x \end{array} \right. \end{array}$$

Since assertion A_2 implies $\neg b_U \vee \neg b_V = \text{T}$, atomic system $\{E_1, E_2\}$ entirely determines the pair (x, y) as its output. A correct abstraction for (86) is thus

$$\{E_1, E_2\} \rightarrow (x, y)$$

Consider next the following variation of (85), in which E_1 has been modified so that x is now a state:

$$\begin{array}{ll} E_1: & x^\bullet = f(x, y) \\ A_2: & \textbf{assert } b_U \wedge b_V = \text{F} \\ E_{21}: & b_U = [U(x) > 0] \\ E_{22}: & b_V = [V(y) > 0] \\ E_{23}: & \textbf{if } b_U \textbf{ then } 0 = V(y) \textbf{ else } 0 = U(x) \\ E_{24}: & \textbf{if } b_V \textbf{ then } 0 = U(x) \textbf{ else } 0 = V(y) \end{array} \tag{87}$$

This is no longer of index 0, so we must shift the complementarity condition. Shifting its causality analysis yields the following, where we have eliminated the unnecessary $y^\bullet$ and taken into account that x is assumed consistent (and thus is no longer occurring in the causality analysis):

$$\begin{array}{ll} A_2: & \textbf{assert } b_U \wedge b_V = \text{F} \\ A_2^\bullet: & \textbf{assert } b_U^\bullet \wedge b_V^\bullet = \text{F} \\ E_2: & \left\{ \begin{array}{l} x \rightarrow E_2 \rightarrow (b_U, b_V) \\ b_U \rightarrow E_{23}, \textbf{if } b_U \textbf{ then } b_U \rightarrow F_{23} \rightarrow y \\ b_V \rightarrow E_{24}, \textbf{if } b_V \textbf{ else } \ b_V \rightarrow F_{23} \rightarrow y \end{array} \right. \\ E_2^\bullet: & \left\{ \begin{array}{l} x^\bullet \rightarrow E_2^\bullet \rightarrow (b_U^\bullet, b_V^\bullet) \\ b_U^\bullet \rightarrow E_{23}^\bullet, \textbf{if } b_U^\bullet \textbf{ else } \ b_U^\bullet \rightarrow F_{24}^\bullet \rightarrow x^\bullet \\ b_V^\bullet \rightarrow E_{24}^\bullet, \textbf{if } b_V^\bullet \textbf{ then } b_V^\bullet \rightarrow F_{24}^\bullet \rightarrow x^\bullet \end{array} \right. \end{array}$$

Taking E_1 into account yields:

$$\begin{array}{ll} A_2: & \textbf{assert } b_U \wedge b_V = \text{F} \\ A_2^\bullet: & \textbf{assert } b_U^\bullet \wedge b_V^\bullet = \text{F} \end{array} \tag{88}$$

atom $\{E_1, E_2, E_2^\bullet\}$:

$$\left\{ \begin{array}{l} x \rightarrow E_2 \rightarrow (b_U, b_V) \\ x^\bullet \rightarrow E_2^\bullet \rightarrow (b_U^\bullet, b_V^\bullet) \\ b_U \rightarrow E_{23}, \textbf{if } b_U \textbf{ then } b_U \rightarrow F_{23} \rightarrow y \rightarrow E_1 \rightarrow x^\bullet \\ b_V \rightarrow E_{24}, \textbf{if } b_V \textbf{ else } \ b_V \rightarrow F_{23} \rightarrow y \rightarrow E_1 \rightarrow x^\bullet \\ b_U^\bullet \rightarrow E_{23}^\bullet, \textbf{if } b_U^\bullet \textbf{ else } \ b_U^\bullet \rightarrow F_{24}^\bullet \rightarrow x^\bullet \rightarrow E_1 \rightarrow y \\ b_V^\bullet \rightarrow E_{24}^\bullet, \textbf{if } b_V^\bullet \textbf{ then } b_V^\bullet \rightarrow F_{24}^\bullet \rightarrow x^\bullet \rightarrow E_1 \rightarrow y \end{array} \right.$$

which determinates $(y, x^\bullet)$ as its ouput if the following condition holds:

$$b_U \vee \neg b_V \vee \neg b_U^\bullet \vee b_V^\bullet \quad = \quad \text{T} \tag{89}$$

Unfortunately, assuming assertion $A_2 \wedge A_2^\bullet$ still does not make (89) a tautology. Extending the array (88) will not solve the problem. This suggests that our assertion A_2 may not be strong enough. Let us strenghten A_2 as follows:

$$A_2 \quad : \quad \left\{ \begin{array}{l} \textbf{assert } b_U \wedge b_V = \text{F} \\ \textbf{assert } [b_U = b_U^\bullet] \vee [b_V = b_V^\bullet] = \text{T} \end{array} \right. \tag{90}$$

which amounts to assuming that the two unilateral constraints do not change their status saturated/unsaturated simultaneously. Then, assuming (90) makes (89) a tautology, showing that the atom $(A_2, \{E_1, E_2\})$ has index 1 under (90).

V. Index reduction and nonstandard semantics

So far we have developed a theory of index for dAE with associated graph based algorithms. A key step in our agenda of developing a theory of index for hybrid DAE systems is to cast smooth DAE systems into the world of dAEs and see what our dAE index theory gives. Our aim is to show that the two ways of defining an index for this class of systems actually coincide. Casting smooth DAE systems into the world of dAEs is performed by interpreting DAE using nonstandard analysis [23], [9], [17], [10], [6].

A. Nonstandard analysis for the engineer

Here we collect the minimum needed to follow the subsequent developments (excluding the proofs). Corresponding details are found in the Appendix B. We will use ${}^\star\mathbb{N}$ and ${}^\star\mathbb{R}$, which are the nonstandard extensions of $\mathbb{N}$ and $\mathbb{R}$, respectively. ${}^\star\mathbb{N}$ is a totally ordered set obtained by extending $\mathbb{N}$ with infinite integers, greater than any standard finite integer. ${}^\star\mathbb{R}$ is obtained by extending $\mathbb{R}$ in two ways: 1) by adding *infinitesimals*, which are nonzero elements ϵ closer to zero than any standard (ordinary) nonzero real, and 2) infinite reals, greater than any standard finite real; to every finite real x we then add its nonstandard neighbours $x+\epsilon$ where ϵ is any infinitesimal. ${}^\star\mathbb{R}$ is totally ordered, and usual operations ($+$, $\times$, etc.) and relations on reals extend to it. For any finite positive real t and any finite nonstandard positive real ∂, there exists $n \in {}^\star\mathbb{N}$ such that $n \times \partial \leq t < (n+1) \times \partial$. This applies in particular when ∂ is infinitesimal, in which case n must be an infinite nonstandard integer.

This suggests to consider the following set as a time domain:

$$\mathbb{T} \quad = \quad \{k\partial \mid k \in {}^\star\mathbb{Z}\}$$

where ∂ is an infinitesimal time step and ${}^{\star}\mathbb{Z}$ is the set of nonstandard positive or negative integers. Elements of $\mathbb{T}$ will be denoted by the symbol τ. Time domain $\mathbb{T}$ is totally ordered and ranges from minus infinity to plus infinity. It is both

"continuous": any finite real standard number t has an element $\tau \in \mathbb{T}$ such that $|t-\tau|$ is infinitesimal; and
"discrete": each $\tau = n\times\partial \in \mathbb{T}$ has a unique immediate predecessor ${}^\bullet\tau =_{\rm def} (n-1)\times\partial$ and immediate successor $\tau^\bullet =_{\rm def} (n+1)\times\partial$.

B. Nonstandard semantics of DAEs

The *nonstandard semantics* of a DAE is obtained by applying the following substitution rules:

$$\begin{array}{lcl} \dot{x} & \leftarrow & \frac{1}{\partial}(x^\bullet - x) \\ \ddot{x} & \leftarrow & \frac{1}{\partial^2}(x^{\bullet 2} - 2\,x^\bullet + x) \\ x^{(3)} & \leftarrow & \frac{1}{\partial^3}(x^{\bullet 3} - 3\,x^{\bullet 2} + 3\,x^\bullet - x) \\ x^{(4)} & \leftarrow & \ldots \end{array} \tag{91}$$

where $x^\bullet, x^{\bullet 2}, x^{\bullet 3}, \ldots$ are defined as follows (we will need both the *pre* operator ${}^\bullet x$ and *post* operator $x^\bullet$ in the sequel):

$$\begin{array}{llllll} {}^\bullet\tau & =_{\rm def} & \tau - \partial & , \quad \tau^\bullet & =_{\rm def} & \tau + \partial \\ {}^\bullet x_\tau & =_{\rm def} & x_{{}^\bullet\tau} & , \quad x^\bullet_\tau & =_{\rm def} & x_{\tau^\bullet} \\ {}^{\bullet m}x & =_{\rm def} & {}^\bullet({}^{\bullet m-1}x) & , \quad x^{\bullet m} & =_{\rm def} & (x^{\bullet m-1})^\bullet \end{array} \tag{92}$$

Applying (91) to the pendulum example (15) yields:

$$\begin{array}{lcl} x^\bullet & = & x + \partial \times u \\ u^\bullet & = & u + \partial \times Tx \\ y^\bullet & = & y + \partial \times v \\ v^\bullet & = & v + \partial \times (Ty - g) \\ L^2 & = & x^2 + y^2 \end{array} \tag{93}$$

C. The two notions of index coincide

We now establish a link between index reduction for DAE and index reduction for dAE.

Return to the pendulum example. Highest degree shifted variables are $x^\bullet, u^\bullet, y^\bullet, v^\bullet, T$. Corresponding Jacobian is singular, thus the difference degree is strictly positive. Forward shifting the last equation two times yields:

$$\begin{array}{lcll} x^\bullet & = & x + \partial \times u & (i1) \\ u^\bullet & = & u + \partial \times Tx & (i2) \\ y^\bullet & = & y + \partial \times v & (ii1) \\ v^\bullet & = & v + \partial \times (Ty - g) & (ii2) \\ L^2 & = & x^2 + y^2 & (iii) \\ L^2 & = & (x^\bullet)^2 + (y^\bullet)^2 & (iv) \\ L^2 & = & (x^{\bullet 2})^2 + (y^{\bullet 2})^2 & (v) \end{array} \tag{94}$$

Substituting, in (iv, v), $x^\bullet$ and $y^\bullet$ by using $(i1)$ and $(ii1)$, and reorganizing the result yields:

$$\begin{array}{lcll} x^\bullet & = & x + \partial \times u & (i1) \\ u^\bullet & = & u + \partial \times Tx & (i2) \\ y^\bullet & = & y + \partial \times v & (ii1) \\ v^\bullet & = & v + \partial \times (Ty - g) & (ii2) \\ L^2 & = & (x^\bullet + \partial \times u^\bullet)^2 + (y^\bullet + \partial \times v^\bullet)^2 & (v) \\ & & - - - - - - - - - - & \\ L^2 & = & x^2 + y^2 & (iii) \\ L^2 & = & (x + \partial \times u)^2 + (y + \partial \times v)^2 & (iv) \end{array} \tag{95}$$

The first group of equations has structurally nonsingular Jacobian with respect to $x^\bullet, u^\bullet, y^\bullet, v^\bullet, T$, and is thus a dAE of index 0. dAE system (93) has thus index 2. All of this suggests that we have the following "pseudo-equation":

$$\text{index(NS(DAE))} \quad \stackrel{\text{structurally}}{=} \quad \text{index(DAE)}$$

where NS(DAE) denotes the nonstandard interpretation of DAE, seen as a dAE. Let us formalize this. In the following, n denotes the dimension of variable v.

Definition 9: Say that the following property holds:

$$x \to \exists w.F(x,v,w) = 0 \text{ structurally defines } v \tag{96}$$

if the pair $(F'_v(x,v,w), F'_w(x,v,w))$ *of Jacobians is structurally nonsingular, see Definition 3 and Lemma 2.*

Informally, if we perturb F while respecting the independence of a given equation with respect to a given variable, the map $x \to \exists w.F(x,v,w) = 0$ defines v almost everywhere. Next, consider

- the differential array $F_k(x,v,w)$ associated with DAE $F(x,\dot{x}) = 0$ following (7), and
- the difference array $F^{\rm NS}_k(x,v,w)$ associated with $F^{\rm NS}(x,x^\bullet) = 0$ following (27), where $F^{\rm NS}(x,x^\bullet)$ is the nonstandard semantics of $F(x,\dot{x})$.

Let us relate these two arrays. In doing so we use different notations for the variables involved in F, written x, v, w, and the variables involved in $F^{\rm NS}$, written $x, \bar{v}, \bar{w}$—we do not need to distinguish the x-variable for the two cases. Coordinates of w are $w_2 \ldots w_k$ and similarly for $\bar{w}$. We have, with reference to (7) and (27):

$$\begin{array}{rcl} F^{\rm NS}_k(x,\bar{v},\bar{w}) & = & \left[\begin{array}{c} F^{(0)}_{\rm NS}(x,\bar{v},\bar{w}) \\ F^{(1)}_{\rm NS}(x,\bar{v},\bar{w}) \\ \vdots \\ F^{(k)}_{\rm NS}(x,\bar{v},\bar{w}) \end{array}\right] \\ F_k(x,v,w) & = & \left[\begin{array}{c} F^{(0)}(x,v,w) \\ F^{(1)}(x,v,w) \\ \vdots \\ F^{(k)}(x,v,w) \end{array}\right] \end{array} \tag{97}$$

Using (91), we get that $(x,\bar{v},\bar{w})$ and (x,v,w) are related by the following formulas, where

$$\binom{i}{m} = \frac{i!}{m!(i-m)!}$$

denote the binomial coefficients:

$$\begin{array}{rcl} v & = & \frac{1}{\partial}(\bar{v} - x) \\ w_2 & = & \frac{1}{\partial^2}(\bar{w}_2 - 2\,\bar{v} + x) \\ w_3 & = & \frac{1}{\partial^3}(\bar{w}_3 - 3\,\bar{w}_2 + 3\,\bar{v} - x) \\ & \vdots & \\ w_i & = & \frac{1}{\partial^i}\sum_{j=0}^{i}(-1)^{i-j}\binom{i}{j}\bar{w}_j \end{array} \tag{98}$$

with the convention that $\bar{w}_1 = \bar{v}$ and $\bar{w}_0 = x$. Then,

$$\begin{array}{rcl} F^{(0)}(x,v,w) & =_{\text{def}} & F(x,v) = F(x, \frac{1}{\partial}(\bar{v} - x)) \\ & =_{\text{def}} & F^{(0)}_{\text{NS}}(x,\bar{v},\bar{w})\,. \end{array} \tag{99}$$

$F^{(i)}(x,v,w)$ is obtained by differentiating i-times $F(x,\dot{x})$ with respect to time and then replacing the successive derivatives of x by $v, w_2, w_3, \ldots, w_i$. Similarly, $F^{(i)}_{\text{NS}}(x,\bar{v},\bar{w})$ is obtained by shifting forward i-times $F(x, \frac{1}{\partial}(x^\bullet - x))$ and then replacing the succesive forward shifts of x by $\bar{v}, \bar{w}_2, \bar{w}_3, \ldots, \bar{w}_i$. Hence, $F^{\text{NS}}_k(x,\bar{v},\bar{w})$ and $F_k(x,v,w)$ are related by

$$\begin{array}{rcl} F^{(0)}(x,v,w) & = & F^{(0)}_{\text{NS}}(x,\bar{v},\bar{w}) \\ F^{(1)}(x,v,w) & = & \frac{1}{\partial}\left(F^{(1)}_{\text{NS}}(x,\bar{v},\bar{w}) - F^{(0)}_{\text{NS}}(x,\bar{v},\bar{w})\right) \\ F^{(2)}(x,v,w) & = & \frac{1}{\partial^2}\left(F^{(2)}_{\text{NS}}(x,\bar{v},\bar{w}) - 2F^{(1)}_{\text{NS}}(x,\bar{v},\bar{w})\right. \\ & & \left. \quad + F^{(0)}_{\text{NS}}(x,\bar{v},\bar{w})\right) \\ & \vdots & \\ F^{(i)}(x,v,w) & = & \frac{1}{\partial^i}\sum_{j=0}^{i}(-1)^{i-j}\binom{i}{j}F^{(j)}_{\text{NS}}(x,\bar{v},\bar{w}) \end{array} \tag{100}$$

Accordingly, if we define the square lower triangular matrix matrix $L(p)$ by

$$\text{for } j \leq i: \quad L(p)_{ij} = \frac{(-1)^{i-j}}{\partial^i}\binom{i}{j}$$

and set

$$L \;=\; L(k) \otimes I_n \tag{101}$$

where $\otimes$ denotes the Kronecker product and I_n is the $n\times n$-identity matrix,[6] the following holds:

Lemma 3: The matrix L defined in (101) is block lower triangular (BLT), invertible and with BLT inverse. Furthermore,

$$\begin{bmatrix} v \\ w_2 \\ w_3 \\ \vdots \\ w_k \end{bmatrix} = L \begin{bmatrix} \bar{v} \\ \bar{w}_2 \\ \bar{w}_3 \\ \vdots \\ \bar{w}_k \end{bmatrix} + \begin{bmatrix} -1/\partial \\ 1/\partial^2 \\ -1/\partial^3 \\ \vdots \end{bmatrix} x \tag{102}$$

$$\begin{bmatrix} F^{(1)} \\ F^{(2)} \\ F^{(3)} \\ \vdots \\ F^{(k)} \end{bmatrix} = L \begin{bmatrix} F^{(1)}_{\text{NS}} \\ F^{(2)}_{\text{NS}} \\ F^{(3)}_{\text{NS}} \\ \vdots \\ F^{(k)}_{\text{NS}} \end{bmatrix} + \begin{bmatrix} -1/\partial \\ 1/\partial^2 \\ -1/\partial^3 \\ \vdots \\ \end{bmatrix} F^{(0)}_{\text{NS}} \tag{103}$$

On the other hand, the map

$$(x,v,w) \;\to\; F_k(x,v,w)$$

factorizes as the following composition of maps:

$$(x,v,w) \overset{(102)}{\to} (x,\bar{v},\bar{w}) \;\to\; F^{\text{NS}}_k(x,\bar{v},\bar{w}) \overset{(103)}{\to} F_k(x,v,w)$$

Using the formula giving the Jacobian of the composition of maps, we get

$$\nabla_{x,v,w}F_k \;=\; \begin{bmatrix} I_n & 0 \\ 0 & L \end{bmatrix} \nabla_{x,\bar{v},\bar{w}}F^{\text{NS}}_k \begin{bmatrix} I_n & 0 \\ 0 & L^{-1} \end{bmatrix}$$

where the notation $\nabla_{x,v,w}f$ stands for the Jacobian of f with respect to the triple (x,v,w).

Theorem 4: The following two properties are equivalent:

$$x \to \exists w.F_k(x,v,w) = 0 \quad \textit{structurally defines } v \tag{104}$$

$$x \to \exists w.F^{\text{NS}}_k(x,\bar{v},\bar{w}) = 0 \quad \textit{structurally defines } \bar{v} \tag{105}$$

Proof: By Lemma 2, property (104) holds if and only if

$$\begin{bmatrix} A & C \end{bmatrix} \;=_{\text{def}}\; \begin{bmatrix} \nabla_v F_k & \nabla_w F_k \end{bmatrix}$$

satisfies (21), for some two permutation matrices P and Q. It turns out that, if L is block-lower triangular, $[A\;C]$ satisfies (21) if and only if so does $L[A\;C]L^{-1}$. ∎

Using an implicit Euler scheme: So far the substitution rules (91) correspond to applying an explicit Euler scheme with an infinitesimal step ∂, see for example (93). What happens if we use an implicit Euler scheme instead, meaning that the nonstandard semantics of a DAE becomes

$$F(\dot{x},x) = 0 \quad \leftarrow \quad F(\tfrac{1}{\partial}(x^\bullet - x), x^\bullet) = 0 \tag{106}$$

For example, ODE $\dot{x} = f(x)$ yields $x^\bullet = x + \partial \times f(x^\bullet)$. Using (106) instead of (91) amounts to replacing, in (100), x by $x^\bullet$ while leaving the rest unchanged. The same substitution holds in (102) while (103) remains unchanged. The substitution $x \leftarrow x^\bullet$ in (102) does not affect the block-triangular matrix L. The bottom line is that Lemma 3 and thus also Theorem 4 still hold if an implicit Euler scheme is used.

[6] The Kronecker product $A \otimes B$ is the block matrix whose ijth block is $a_{ij}B$.

VI. Hybrid DAE

In this section we apply our previous results fo Hybrid DAE systems. We first describe the class of hybrid DAE systems we consider. Then, we describe a set of primitives that is sufficient to specify hybrid DAE systems. Finally, we study their index.

A. Trajectories

We collect in a preamble the needed notations regarding trajectories of hybrid systems. The reader is referred to Figure 1.

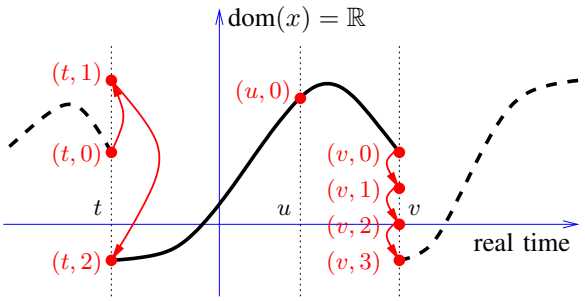


Figure 1. Trajectory of a hybrid system.

This figure shows the trajectory of a variable x of type real. We will be interested in a restricted class of hybrid systems, in which:

- Variables of type real possess trajectories such as in Figure 1, meaning that
  - instants of discontinuity are isolated (t and v in the figure); at t the dashed trajectory hits the red bullet labeled with $(t,0)$, and then jumps to the red bullet labeled with $(t,1)$ and then $(t,2)$; similarly, at time v, the trajectory has its second discontinuity, by successively hitting $(v,0),(v,1),(v,2)$, and $(v,3)$.
  - the trajectory is continuous in the closed interval $[t,v]$, with a beginning depicted by the red bullet labeled with $(t,2)$ and an end depicted by the red bullet labeled with $(v,0)$. During the open interval (t,v), points on the trajectory are labeled with time index $(u,0)$, indicating that the associated date of occurrence is u (thus, the second field 0 of the date serves nothing).
- The same holds for all variables having a domain equipped with a topology. In particular, for discrete domains (such as Booleans or Integers), the same holds if the domains are equipped with the discrete topology.

Formally, we use as time index the set $\mathbb{S}=\mathbb{R}\times\mathbb{N}$, equipped with the lexicographic order defined by: $(t,k)<(t',k')$ if and only if, either $t<t'$, or $t=t'$ and $k<k'$. Elements of $\mathbb{S}$ are denoted by the symbol s, or explicitly as pairs $s=(t,k)$ whenever needed. By convention:

$$\text{We identify } \mathbb{R} \text{ and the subset } \{s=(t,0)\mid t\in\mathbb{R}\}\subset\mathbb{S}. \tag{107}$$

Time set $\mathbb{S}$ defines the so-called *super-dense time*, see [14], [15]. It allows defining solutions for (110) in which finite (but possibly unbounded) cascades of mode changes can occur.

To every trajectory, we associate a function

$$\mathbb{R}\ni t \quad\to\quad n_t\in\mathbb{N}$$

such that $n_t=0$, except for t belonging to some increasing sequence $T=\{t_k\mid k\in\mathbb{Z}\}$ of instants of $\mathbb{R}$ with $\lim_{k\to\pm\infty}t_k=\pm\infty$. For the trajectory shown in Figure 1, we have $n_t=2$, $n_v=3$, and otherwise n takes the value zero.

For $s\in\mathbb{S}$, define x_s^- and x_s^+ by

$$x_s^- \;=_{\text{def}}\; \begin{cases} \lim_{t'\nearrow t} x_{(t',0)} & \text{if} \quad s=(t,0) \\ x_{(t,k-1)} & \text{if} \quad s=(t,k), k>0 \end{cases} \tag{108}$$

$$x_s^+ \;=_{\text{def}}\; \begin{cases} \lim_{t'\searrow t} x_{(t',0)} & \text{if} \quad s=(t,n_t) \\ x_{(t,k+1)} & \text{if} \quad s=(t,k), k<n_t \end{cases} \tag{109}$$

B. Mode dependent dynamics

In this section we use the guarded equations introduced in Section IV-B and the reader is referred to this section for the notations. The basic form for a hybrid DAE system is the following, where I is some finite set of *modes* and i ranges over I:

$$S \quad : \quad \forall i\in I \begin{cases} b_i = B_i(x^-,x,\dot{x}) \\ \textbf{if } b_i \textbf{ then } 0=F_i(x^-,x,\dot{x}) \end{cases} \tag{110}$$

where I is some finite index set, x denotes a n-tuple of real variables, denote the left- and right-limit of x at instant t, the F_i's are real-valued and smooth, and the B_i's are smooth boolean predicates over the listed variables. The first equation specifies that, in mode i, DAE $F_i=0$ must hold; observe that this is a fixpoint equation. Thus, b_i is the *guard* of mode i and $F_i=0$ its dynamics. If two or more guards overlap, then the conjunction of the corresponding dynamics must hold. If the "or" of all guards is not the constant "true", then S is incompletely specified.

Although we do not require that the different modes satisfy $\bigvee_{i\in I} b_i = \text{T}$ and $b_i\wedge b_j=\text{F}$ for any two $i\neq j$, this will generally be the case in practice.

The generic form (110) does not allow for nested guards, e.g., of the form **if** b' **then** [**if** b **then** E]. Flattening can be applied to bring nested guards to the form (110). Flattening, however, is known not to be desirable from the point of view of computational complexity, as it may typically give raise to combinatorial explosion. Nested guards are thus desirable to include but are left for future work.

Examples: We discuss here a few examples, showing the flexibility of generic form (110):

A first example is obtained by considering a DAE system with unilateral constraint:

$$0 \quad\le\quad F(x,\dot{x}) \tag{111}$$

where x and F are as above. Equation (111) rewrites

$$\begin{cases} b=[\,0\ge F(x,\dot{x})\,] \\ \textbf{if } b \textbf{ then } 0=F(x,\dot{x}) \end{cases}$$

This system has two modes. In the first mode, $b=\text{T}$ expresses that unilateral constraint (111) is *active*. In the second mode corresponding to the **else** alternative, the constraint is trivial,

expressing that $(x, \dot{x})$ is unconstrained when the inequality is strict in (111).

So-called DAE systems with a *complementarity conditions* are a second example:

$$\begin{array}{r} U(x)\geq 0 \text{ and } V(y)\geq 0 \text{ and } U(x)V(y) = 0 \\ F(x,y) = 0 \end{array} \tag{112}$$

where we assume that the second equation ensures that the system (112) is nonsingular. The first equation of (112)—the complementarity condition—is often written in a compact form as

$$0 \leq U(x) \perp V(y) \geq 0 \tag{113}$$

Such systems are encountered, e.g., in electric circuits with perfect diodes. System (112) can be given the generic form (110) in the following way:

$$\begin{array}{l} 0 = F(x,y) \\ b_U = [U(x){>}0] \\ b_V = [V(x){>}0] \\ \textbf{if } b_U \textbf{ else } [0{=}U(x)] \\ \textbf{if } b_V \textbf{ else } [0{=}V(y)] \\ \textbf{if } b_U \textbf{ then } [0{=}V(y)] \\ \textbf{if } b_V \textbf{ then } [0{=}U(x)] \end{array} \tag{114}$$

A third example is the zero-crossing example (75) of Section IV-D1, where boolean guard b selects the events of *zero-crossings* of some smooth function $g(x)$. □

We now define what a solution of (110) is. Let $\mathbf{B}$ denote the Boolean domain.

Definition 10: Hybrid DAE (110) *is* solvable *if there exists a pair of functions* $(s, \lambda) \to (\Phi(s, \lambda), \beta(s, \lambda))$, *from* $\mathbb{S}\times\Lambda$ *into* $(\mathbb{R}^n \times \mathbf{B}^I)\cup\{\epsilon\}$, *where* ϵ *is the* undefined *value and* Λ *is some nonempty open set of* $\mathbb{R}^\rho$, *satisfying the following conditions, where* $\beta_i, i{\in}I$ *denote the components of* β*:*

1) *For every* λ, *there exists a function* $\mathbb{R} \ni t \to n_t(\lambda) \in \mathbb{N}$, *such that* $(\Phi((t,k),\lambda), \beta((t,k),\lambda)) = \epsilon$ *if and only if* $k > n_t(\lambda)$. *Furthermore,* $n_t(\lambda) = 0$, *except for* t *belonging to some increasing sequence* $T(\lambda) = \{t_k(\lambda) \mid k{\in}\mathbb{Z}\}$ *of instants of* $\mathbb{R}$ *with* $\lim_{k\to\pm\infty} t_k(\lambda){=}\pm\infty$. *We write* t_k *instead of* $t_k(\lambda)$ *when no confusion results.*
2) *Regarding the modes:*
   - *The function* $t{\to}\beta((t,0),\lambda)$ *is constant over every open interval* (t_k, t_{k+1}), *and, letting* $b \in \mathbf{B}^I$ *be the corresponding value, we have* $\beta((t_k, n_{t_k}), \lambda) = b$ *and* $\beta((t_{k+1}, 0), \lambda) = b$.
3) *Regarding the state* x, *and using* (108), (109) *and* (107)*:*
   - *For every* $\lambda \in \Lambda$ *and every interval* (t_k, t_{k+1}), $t{\to}\Phi(t,\lambda)$ *is a diffeomorphism from* (t_k, t_{k+1}) *into* $\mathbb{R}^n$, *and, if* $\beta_i(t,\lambda) = \mathrm{T}$, *then*

     $$F_i\left(\Phi^-(t,\lambda), \Phi(t,\lambda), \tfrac{d}{dt}\Phi(t,\lambda)\right) = 0$$

     holds for every $t \in (t_k, t_{k+1})$.
   - *For every* $\lambda \in \Lambda$ *and every* $t{\in}T(\lambda)$, *then*

     $$\begin{array}{rcl} \Phi((t,0),\lambda) & = & \Phi^-((t,0),\lambda) \\ \Phi((t,n_t),\lambda) & = & \Phi^+((t,0),\lambda) \end{array}$$

     and, for every $k{>}0$ *and* i *such that* $\beta_i((t,k),\lambda){=}\mathrm{T}$, *then* $x_i =_{\mathrm{def}} \Phi((t,k),\lambda)$ *is a consistent value for* $F_i(x^-, x, \dot{x}) = 0$.
4) *If* $s{\to}(x_s, b_s)$ *satisfies conditions* 1)–3) *above, then it holds that* $(x_s, b_s) = (\Phi(s,\lambda), \beta(s,\lambda))$ *for some* λ.

Some comments are in order regarding Definition 10, with reference to its successive conditions:

1) The function $n_t(\lambda)$ speficies, for instant t, whether the system traverses a cascade of mode changes ($n_t(\lambda) > 0$ being the length of this cascade), or not ($n_t(\lambda) = 0$). Cascades are assumed to be finite (but not necessarily uniformly bounded) and isolated from each other.
2) A mode begins at the last instant of the last seen cascade (where it is reset), and ends at the first instant of the next cascade.
3) The system dynamics at a persistent mode i (meaning that $b_i = \mathrm{T}$ for some positive duration) is F_i.
4) Parameter λ serves to parameterize the solutions by fixing the consistent resets, for each mode.

Our definition of a solution for a hybrid DAE system generalizes the classical definition for (ODE based) hybrid systems. Observe that conditions for existence and/or uniqueness of solutions are delicate, particularly so because we have ruled out Zeno behaviors.

The theory of DAE differentiation index recalled in Section II-A deeply relies on differentiability, so it does not apply as such to (110). In contrast, the notion of difference index for dAE does not require differentiability. To circumvent the lack of differentiability of (110), we move to its nonstandard semantics. As a matter of fact, the nonstandard semantics of a DAE hybrid system is simple and clean defining.

C. Nonstandard hybrid DAE index

Using (91) and (92), the nonstandard semantics of (110) is the following constraint:

$$H_{\mathrm{NS}}({}^\bullet x, x, x') = \left[\begin{array}{l} b_i = B_i({}^\bullet x, x, x') \\ \textbf{if } b_i \textbf{ then } 0 = F_i({}^\bullet x, x, x') \end{array}\right] \tag{115}$$

$$\text{where } x'_\tau = \frac{x_{\tau+\partial} - x_\tau}{\partial} \text{ i.e., } x' = \frac{x^\bullet - x}{\partial} \tag{116}$$

That is, we substitute, in S, the left limit x^- by the previous value ${}^\bullet x$ defined in (92) and the derivative $\dot{x}$ by its nonstandard version defined in (91). In the following reasoning, the concepts and notations of Appendix B are used.

Theorem 5: Assume that hybrid DAE (110) *is solvable. Then, every solution of* (110) *is the standardisation of some solution of* (115).

Proof: See Appendix A4. The proof relies on the material of Appendix B. ■

We now study the difference index of dAE (115). Accordingly, using (116) we regard F as a function of the tuple $({}^\bullet x, x, x^\bullet)$. To simplify the notations, when no confusion can result, we write F for short instead of $F({}^\bullet x, x, x^\bullet)$ and similarly for B. With this convention, we have

$$H_{\rm NS}^{\bullet k}(^\bullet x, x, v, w) = \left[\begin{array}{l} b_i^{\bullet k} = B_i^{\bullet k}(^\bullet x, x, x') \\ \textbf{if } b_i^{\bullet k} \textbf{ then } 0 = F_i^{\bullet k}(^\bullet x, x, x') \end{array}\right] \quad (117)$$

where

$$v = x^\bullet \quad \text{and} \quad w = (x^{\bullet 2}, \ldots, x^{\bullet k+1}) \quad (118)$$

and x' is defined by (116). Using (117), the difference array of dAE (115) is

$$H_k^{\rm NS}(^\bullet x, x, v, w) =_{\rm def} \left[\begin{array}{l} H_{\rm NS}(^\bullet x, x, v, w) \\ H_{\rm NS}^{\bullet}(^\bullet x, x, v, w) \\ \vdots \\ H_{\rm NS}^{\bullet k}(^\bullet x, x, v, w) \end{array}\right] \quad (119)$$

Array (119) is amenable to Definition 8 and Theorem 3, which yields a guarded Pantelides algorithm for computing the causality and performing index reduction. As announced in the introduction, the causality is dynamically defined, that is, it depends on the sequence of modes traversed by the system at the considered successive instants $t, \ldots, t^{\bullet k}$.

VII. Analyzing some examples

In this section, we analyze some examples. Most of them were already investigated in Section IV-D under their dAE variation.

A. Zero-crossing

We reconsider the example (75) of Section IV-D:

$$\begin{array}{l} \textbf{if } b \textbf{ then } x = h(x^-) \textbf{ else } \dot{x} = f(x) \\ b = \text{zero-crossing of } g(x) \end{array} \quad (120)$$

where boolean guard b selects the events of *zero-crossings* of some smooth function $g(x)$. Our first task is to make the semantics of (120) precise, particularly at and around the zero-crossing. We do this now by defining different versions of the the nonstandard semantics of (120). The different versions for the nonstandard semantics are:

$$\left\{\begin{array}{l} \textbf{if } b \textbf{ then } x = h(^\bullet x) \textbf{ else } x' = f(x) \\ b = Q(x, x^\bullet) \end{array}\right. \quad (121)$$

$$\left\{\begin{array}{l} \textbf{if } b \textbf{ then } x = h(^\bullet x) \textbf{ else } x' = f(x) \\ b = Q(^\bullet x, x) \end{array}\right. \quad (122)$$

$$\left\{\begin{array}{l} \textbf{if } b \textbf{ then } x^\bullet = h(x) \textbf{ else } x' = f(x) \\ b = Q(^\bullet x, x) \end{array}\right. \quad (123)$$

where

$$\begin{array}{rcl} x' & =_{\rm def} & \frac{x^\bullet - x}{\partial} \quad \text{and} \\ Q(z, x) & =_{\rm def} & [g(z) \leq 0] \wedge [g(x) > 0] \end{array} \quad (124)$$

Semantics (121) is closest to (120). Semantics (123) introduces micro-delays to ease causality—to this end, the predicate defining guard b has been shifted backward and reset has been shifted forward. Semantics (122) lies in between the former two. Semantics (121), (122), and (123) have the form (76), (77), and (78), respectively. The analysis developed in Section IV-D1 shows that the only suitable semantics for zero-crossing (120) is (123). The discussion at the end of Section IV-D1 carries over to semantics (123).

Conclusions regarding this example: A first generic conclusion is that we must perform mode-dependent causality analyses by not abstracting predicate evaluation. This was developed in Section IV-B.

The second, specific, conclusion is that we should be careful in defining the semantics of events (such as zero-crossings). Not putting appropriate micro-delays in the nonstandard semantics results in semantics with causality circuits that cannot get repaired by using higher order difference arrays, evidencing that those semantics have infinite index.

A generic form of zero-crossing: Equation (75) specifies a system in which the state is reset each time some zero-crossing event occurs. The generic case is when the zero-crossing event triggers the move from one arbitrary mode to another arbitrary mode. A possible continuous-time form for this system is the following, where the two modes 0 and 1 are encoded by the value of the piecewise constant variable ξ (its dynamics is $\dot{\xi} = 0$), which takes first the value 0 and then the value 1:

$$\left\{\begin{array}{l} \textbf{init } \xi = 0 \\ \textbf{if } \xi = 0 \textbf{ then } b = \textbf{zero-crossing of } g(x) \\ \textbf{if } b \textbf{ then } \xi = 1 \textbf{ else } \dot{\xi} = 0 \\ \textbf{if } \xi = 0 \textbf{ then } F_0(x, \dot{x}) = 0 \\ \qquad \textbf{else if } \xi = 1 \textbf{ then } F_1(x, \dot{x}) = 0 \end{array}\right. \quad (125)$$

System (125) can be put in the generic form (110):

$$\left\{\begin{array}{l} \textbf{init } \xi = 0 \\ \textbf{if } b \textbf{ then } \xi = 1 \textbf{ else } \dot{\xi} = 0 \\ \textbf{if } \xi{=}0 \textbf{ then } \left[\begin{array}{l} F_0(x, \dot{x}) = 0 \\ b = \textbf{zero-crossing of } g(x) \end{array}\right] \\ \textbf{if } \xi{=}1 \textbf{ then } \left[\begin{array}{l} F_1(x, \dot{x}) = 0 \end{array}\right] \end{array}\right. \quad (126)$$

A similar analysis as for the simpler zero-crossing example shows that the appropriate nonstandard semantics for (126) is the following, where $Q(z, x)$ is as in (124):

$$\left\{\begin{array}{l} \textbf{init } \xi = 0 \\ \textbf{if } b \textbf{ then } \xi^\bullet = 1 \textbf{ else } \xi' = 0 \\ \textbf{if } [\xi{=}0] \textbf{ then } \left[\begin{array}{l} F_0(x, x'){=}0 \\ b{=}Q(^\bullet x, x) \end{array}\right] \\ \textbf{if } [\xi{=}1] \textbf{ then } \left[\begin{array}{l} F_1(x, x'){=}0 \end{array}\right] \end{array}\right. \quad (127)$$

Depending on the particular form for the two dynamics F_0 and F_1, a consistent causality may be found for (127) exactly as for the simple zero-crossing case. The key point in obtaining this is the additional delay applied to the reset of ξ in the equation **if** b **then** $[\xi^\bullet{=}1]$, which ensures that the relation between x, b, and ξ is not fixpoint but can rather be properly scheduled for its evaluation.

B. Unilateral constraint

In this section we reconsider the example of unilateral constraint introduced in Section IV-D, albeit in continuous

time:

$$\begin{cases} \dot{x} = f(x,u) \\ 0 \le g(x) \end{cases}$$

Here is a word for word translation to nonstandard semantics:

$$\begin{cases} x' = f(x,u) \\ b = [g(x) \le 0] \\ \textbf{if } b \textbf{ then } g(x) = 0 \end{cases} \tag{128}$$

where we recall that x' is the nonstandard translation of $\dot{x}$ given in (124). This formulation has the form studied in Section III-E3.

C. Complementarity condition

Replacing, in example (114), the derivatives $\dot{x}$ and $\dot{y}$ by their nonstandard form yields a system of the form studied in Section IV-D3.

D. Circuit breaker

Figure 2 shows a simple circuit breaker.

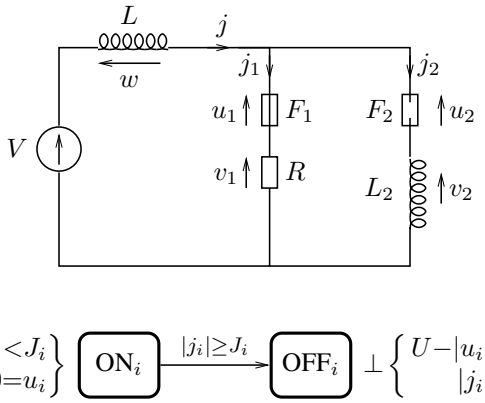


Figure 2. A simple circuit breaker. Top: the circuit. Bottom: the mode automaton for each fuse $i = 1, 2$. For the ON mode, the current must stay below a threshold J_i, while in the OFF mode, the complementarity condition shown holds. We assume that initial conditions are: fuse F_1 is ON, fuse F_2 is OFF with $j_2 = 0$ (no current), and J_1 is small enough so that fuse F_1 eventually switches to OFF.

Mode-independent equations for this circuit are:

$$\begin{array}{rcl} 0 & = & j_1 + j_2 - j \\ 0 & = & u_1 + v_1 + w - V \\ 0 & = & u_1 + v_1 - u_2 - v_2 \\ 0 & = & v_1 - Rj_1 \\ 0 & = & Lj' - w \\ 0 & = & L_2 j_2' - v_2 \end{array} \tag{129}$$

whereas the equations of the two fuses are mode-dependent:

$$\begin{array}{rcl} \text{while in ON}_i & : & \begin{cases} |j_i| < J_i \\ 0 = u_i \end{cases} \\ \text{event ON2OFF}_i & : & |j_i| \ge J_i \\ \text{while in OFF}_i & : & \perp \begin{cases} U - |u_i| \ge 0 \\ |j_i| \ge 0 \end{cases} \end{array} \tag{130}$$

We begin with an informal review of this model, by referring to the classical notion of DAE differentiation index, which we apply in each mode separately—we know this does not give the right notion of index for the whole system.

Focus on the complementarity condition in OFF_i mode. It involves u_i, which is not a state, and j_i, which is a state. Hence, the two saturated constraints yield dynamics of index 0 and 1 in the two complementary modes of OFF_1, respectively. As a result, the system moves from index 1 to index 0-or-1 dynamics for the fuse F_2 as a consequence of the switching of F_1 to OFF mode.

Let us detail the nonstandard semantics, by reusing our approach (125)–(127) to handle the transition $\text{ON}_i \to \text{OFF}_i$, which is by a zero-crossing:

$$\begin{cases} \textbf{init } \xi = 0 \\ \textbf{if } b \textbf{ then } \xi^\bullet = 1 \textbf{ else } \xi' = 0 \\ \textbf{if } [\xi{=}0] \textbf{ then } \begin{bmatrix} E_{\text{ON}_1} \\ b{=}Q(^\bullet x, x) \end{bmatrix} \\ \textbf{if } [\xi{=}1] \textbf{ then } \begin{bmatrix} E_{\text{OFF}_1} \end{bmatrix} \end{cases} \tag{131}$$

where E_{ON_1} and E_{OFF_1} are the nonstandard dynamics of the circuit in modes ON_1 and OFF_1, respectively:

$$E_{\text{ON}_1} : \begin{cases} 0 = j_1 + j_2 - j \\ 0 = u_1 + v_1 + w - V \\ 0 = u_1 + v_1 - u_2 - v_2 \\ 0 = v_1 - Rj_1 \\ 0 = Lj' - w \\ 0 = L_2 j_2' - v_2 \\ 0 = u_1 \\ 0 \le U - |u_2| \perp |j_2| \ge 0 \end{cases}$$

$$E_{\text{OFF}_1} : \begin{cases} 0 = j_1 + j_2 - j \\ 0 = u_1 + v_1 + w - V \\ 0 = u_1 + v_1 - u_2 - v_2 \\ 0 = v_1 - Rj_1 \\ 0 = Lj' - w \\ 0 = L_2 j_2' - v_2 \\ 0 \le U - |u_1| \perp |j_1| \ge 0 \\ 0 \le U - |u_2| \perp |j_2| \ge 0 \end{cases}$$

Only the last two equations differ (they are shown in red). The other equations are mode-independent. The two modes involve complementarity conditions. Hence, our first task is to identify the atomic subsets of equations, for the two modes ON_1 and OFF_1. We first consider mode ON_1, which we rewrite in three blocks as follows:

$$E_{\text{ON}_1} : \begin{cases} E_1 : & 0 = [j_1]_{\text{out}} + j_2 - j \\ E_2 : & 0 = [v_1]_{\text{out}} - Rj_1 \\ & - - - - - - - - - - \\ E_{31} : & 0 = L_2 j_2' - v_2 \\ E_{32} : & 0 = u_1 + v_1 + w - V \\ E_{33} : & 0 = u_1 + v_1 - u_2 - v_2 \\ E_{34} : & 0 = u_1 \\ E_{35} : & 0 \le U - |u_2| \perp |j_2| \ge 0 \\ & - - - - - - - - - - \\ E_4 : & 0 = Lj' - w \end{cases}$$

The states are j and j_2. Accordingly, their current value is known and the first block inherits its consistent causality

as shown. The second block collects the equation of the inductance, the two Kirchhoff voltage laws, the constraint on u_1, and the complementarity condition: we make it an atom with state j_2. We handle this atom as the second case of Section IV-D3, showing that it has outputs $v_2, u_1, u_2, w, j_2^\bullet$ as desired. Accordingly, the consistent causality for this mode is the following, where $E_3 =_{\text{def}} \{E_{31}, \ldots, E_{35}\}$:

$$\begin{array}{l}(j_2, j) \rightarrow E_1 \rightarrow j_1\\ j_1 \rightarrow E_2 \rightarrow v_1\\ (j_2, v_1) \rightarrow E_3 \rightarrow (v_2, u_1, u_2, w, j_2^\bullet)\\ w \rightarrow E_4 \rightarrow j^\bullet\end{array}$$

Performing the same for E_{OFF_1} yields:

$$E_{\text{OFF}_1} \quad : \quad \left\{ \begin{array}{ll} E_1: & 0 = [j_1]_{\text{out}} + j_2 - j \\ E_2: & 0 = [v_1]_{\text{out}} - R j_1 \\ & - - - - - - - - - - \\ E_{31}: & 0 = L_2 j_2' - v_2 \\ E_{32}: & 0 = u_1 + v_1 + w - V \\ E_{33}: & 0 = u_1 + v_1 - u_2 - v_2 \\ E_{34}: & 0 \le U - |u_1| \perp |j_1| \ge 0 \\ E_{35}: & 0 \le U - |u_2| \perp |j_2| \ge 0 \\ & - - - - - - - - - - \\ E_4: & 0 = L j' - w \end{array} \right.$$

Focus on E_{34}. When $0 < U - |u_1|$ holds, this complementarity condition enforces $j_1{=}0$, which, together with E_1, makes this system overconstrained, hence singular. This mode is therefore non reachable and E_{34} boils down to

$$E_{34}: \quad U = u_1$$

and, thus, E_{OFF_1} is handled the same way as E_{ON_1}. The resulting scenario for this circuit is then: in a first phase, fuse F_1 is on and fuse F_2 is OFF and open (i.e., such that $j_2{=}0$); the circuit remains so until j_1 crosses threshold J_1; then, fuse F_1 switches to mode OFF and leaks immediately ($j_1{>}0, u_1{=}U$), while the status of fuse F_2 remains open.

Conclusions regarding this example:

- The system has index 1;
- Its causality analysis is mode-dependent.

E. Discussion

The analysis of the above examples shows that our notion of index for DAE hybrid systems is the adequate generalization of the notion of consistent causality for ODE hybrid systems. Our notion of index is particularly useful at handling non-intuitive situations where the index changes at events or cascades thereof. We have shown that it is important not to abstract modes while performing this causality analysis: the right notion of causality must be mode-dependent. Not doing so may result in overly optimistic causality analyses, see the zero-crossing example. In turn, the index itself remains globally defined for the overall system and is a unique integer (not mode-dependent).

VIII. Algorithms

$\mathcal{B}, \mathcal{X}, \mathbb{X}, \mathbb{E}$, are as in Section IV-B. Recall that E denotes equations or atoms. In this section, we provide an encoding of guarded Pantelides graphs as systems of Boolean equations, which we call *Pantelides systems*. Boolean values F, T and encoded as the integers $0, 1$. Then, we consider the following sets of variables with domain $\{0, 1\}$:

$$\left.\begin{array}{r} x(E) \\ \delta(E, x) \end{array}\right\} \text{ where } x \in \mathbb{X}, E \in \mathbb{E} \tag{132}$$

Non-directed and directed branches of the Pantelides graph are encoded as follows:

$$\begin{array}{rll} x(E) = 1 & \text{encodes branch} & E{-}x \\ x(E)\delta(E, x) = 0 & \text{encodes branch} & E{\rightarrow}x \end{array} \tag{133}$$

For the first equation, $x(E)$ takes the value 1 if variable x is involved in equation E and 0 if it is not involved. For the second equation, if x is involved in equation E (hence $x(E) = 1$), the additional term $\delta(E, x)$ is 0 if x results from evaluating E and otherwise (i.e., if it is an input) it takes the value 1. The value of $\delta(E, x)$ does not matter if x is not involved in E. Using (133),

$$(1 - x(E))\, b = 0 \quad \text{encodes branch} \quad \textbf{if } b \textbf{ then } E{-}x.$$

Stating that x is an output of E when x is involved in E and guard $b = \text{T}$ is written:

$$b\, x(E)\, \delta(E, x) = 0$$

In particular, due to condition 3) of Definition 8, we have, for every guard b, the equation B defining it, and every other variable x involved in B,

$$\left\{ \begin{array}{rcl} 0 & = & 1 - x(B) \\ 0 & = & 1 - b(B) \\ 0 & = & 1 - \delta(B, x) \\ 0 & = & \delta(B, b) \end{array} \right. \tag{134}$$

which is equivalent to the single equation

$$[1{-}x(B)] + [1{-}b(B)] + [1{-}\delta(B, x)] + \delta(B, b) = 0$$

In addition to the above encoding of the guarded Pantelides graph using Pantelides systems, additional constraints are for consideration.

Axioms: The following axioms must hold, in order for a Pantelides system to represent a consistent causality: for every $x \in \mathcal{X}$, we have

$$\forall E \in \mathbb{E} : 1 = x(E)\delta(E, {}^\bullet x) \tag{135}$$

$$1 \ge \sum_{E \in \mathbb{E}, b \in \mathcal{B}} b\, x(E)\, (1 - \delta(E, x)) \tag{136}$$

Axiom (136) states that, in any mode (specified by a given guard), any variable cannot be determined by more than one equation. Axiom (135) follows from the fact that a previous variable cannot be an output. □

Goal: The following is the *goal* of index reduction. For every $x \in \mathcal{X}$:

$$1 = \sum_{E \in \mathbb{E}, b \in \mathcal{B}} b\, x(E)\, (1 - \delta(E, x^\bullet)) \tag{137}$$

It expresses that shifted variables must be entirely, not partially, determined. □

Observe that Axioms (135) and (136), as well as goal (137) do not involve guards. Hence (135)–(137) must hold for any reachable configuration of the guards.

IX. Conclusion

To our knowledge, no proper notion of index existed for hybrid DAE systems (multi-mode). By relying on *nonstandard analysis* we were able to propose such an extension.

Nonstandard analysis formalizes differential equations as discrete step transition systems with infinitesimal time basis. The nonstandard semantics of a hybrid DAE system yields a (discrete time) difference Algebraic Equation (dAE), for which the notion of *difference index* can be defined as well—the difference index of a difference Algebraic Equation (dAE) is an easy transposition of the differentiation index, in which forward shift replaces derivation. We proved that the differentiation index of a DAE is equal to the difference index of its nonstandard semantics, *structurally*. "Structurally" means that we are seeking for a "generic equality", i.e., an equality that is valid if the actual parameters defining the system dynamics remain outside some exceptional set. Thanks to this result, we can propose, as a definition of the index of a hybrid DAE system, the index of its nonstandard semantics (which is a dAE system) and this definition is a conservative extension of both the DAE index and the dAE index.

For both DAE and dAE systems, the structural index is computed by using graph based algorithms of Pantelides type. Due the the need for considering modes and their guards, computing the structural index of a dAE system or a hybrid DAE system requires a new *guarded causality analysis,* in which causalities are derived in a mode-dependent way. Whereas causality analyses must become mode dependent, the index in itself remains a global notion. We illustrated all of this on some examples.

The algorithms we proposed avoid listing the different modes explicitly. In our guarded causality analysis, branches of the graph are labeled by a predicate characterizing the set of all states in which the considered branch occurs in the Pantelides graph. Doing this is expected to cope with the combinatorial explosion resulting from composing a large number of multi-mode systems. Also, we encode the search for a guarded causality as an integer programming problem, thus making it possible to solve it in a modular way—our algorithms are modular in that they provide results that can be reused in subsequent contexts, not known in advance. This is not the case for the purely graph based algorithms, which require knowing the entire system before analyzing it.

Acknowledgement: Philippe Chartier is indebted for fruitful discussions.

References

[1] Acary Vincent and Brogliato Bernard, *Numerical Methods for Nonsmooth Dynamical Systems*, ser. Lecture Notes in Applied and Computational Mechanics. Springer-Verlag, 2008, vol. 35.

[2] A. Aubry and P. Chartier, "Pseudo-symplectic Runge-Kutta Methods," *BIT*, vol. 38(3), pp. 229–246, 1998.

[3] A. Benveniste, B. Caillaud, and P. L. Guernic, "Compositionality in dataflow synchronous languages: Specification and distributed code generation," *Inf. Comput.*, vol. 163, no. 1, pp. 125–171, 2000.

[4] A. Benveniste, T. Bourke, B. Caillaud, and M. Pouzet, "Nonstandard semantics of hybrid systems modelers," *J. Comput. Syst. Sci.*, vol. 78, no. 3, pp. 877–910, 2012.

[5] A. Benveniste, P. Caspi, S. A. Edwards, N. Halbwachs, P. L. Guernic, and R. de Simone, "The synchronous languages 12 years later," *Proceedings of the IEEE*, vol. 91, no. 1, pp. 64–83, 2003.

[6] S. Bliudze, "Un cadre formel pour l'étude des systèmes industriels complexes: un exemple basé sur l'infrastructure de l'UMTS," Ph.D. dissertation, Ecole Polytechnique, 2006.

[7] S. CAMPBELL, "A computational method for general higher index nonlinear singular systems of differential equations," *Repr. from Numerical and Applied Mathematics, v. 12, 1989 p 555-560*, 1990.

[8] G. C.W. and L. Petzold, "ODE methods for the solution of differential-algebraic systems," *SIAM J. Numer. Anal.*, vol. 21, pp. 716–728, 1984.

[9] F. Diener and G. Reeb, *Analyse non standard*. Hermann, 1989.

[10] N. C. (ed.), *Nonstandard analysis and its applications*. Cambridge Univ. Press, 1988.

[11] H. Elmqvist, "A structured model language for large continuous systems," 1978, PhD, Lund University.

[12] Gawthrop, Peter J. and Smith, Lorcan P. S., *Metamodelling: bond graphs and dynamic systems*. Prentice Hall, 1996.

[13] Jean Thoma, *Bond graphs: introduction and applications*. Elsevier Science, 1975.

[14] E. Lee and H. Zheng, "Operational semantics of hybrid systems," in *HSCC*, 2005, pp. 25–53.

[15] E. A. Lee and H. Zheng, "Leveraging synchronous language principles for heterogeneous modeling and design of embedded systems," in *EMSOFT*, 2007, pp. 114–123.

[16] Lennart Ljung and Torkel Glad, *Modeling of Dynamic Systems*. Prentice Hall, 1994.

[17] T. Lindstrøm, "An invitation to nonstandard analysis," in *Nonstandard Analysis and its Applications*, N. Cutland, Ed. Cambridge Univ. Press, 1988, pp. 1–105.

[18] S. Mattsson, H. Elmqvist, and M. Otter, "Physical system modeling with Modelica," *Control Engineering Practice*, vol. 6, pp. 501–510, 1998.

[19] C. Pantelides, "The consistent initialization of differential-algebraic systems," *SIAM J. Sci. Stat. Comput.*, vol. 9, no. 2, pp. 213–231, 1988.

[20] Peter Fritzson, *Introduction to Modeling and Simulation of Technical and Physical Systems with Modelica*. IEEE Press, Wiley, 2011.

[21] ——, *Principles of Object-Oriented Modeling and Simulation with Modelica 3.3*. IEEE Press, Wiley, 2014.

[22] L. Petzold, "Differential algebraic equations are not ODEs," *SIAM J. Sci. Stat. Comput.*, vol. 3, pp. 367–384, 1982.

[23] A. Robinson, *Nonstandard Analysis*. Princeton Landmarks in Mathematics, 1996, ISBN 0-691-04490-2.

[24] Stephen L. Campbell and C. William Gear, "The index of general nonlinear DAEs," *Numer. Math.*, vol. 72, pp. 173–196, 1995.

[25] Sven Erik Mattsson and Gustaf Söderlin, "Index reduction in Differential-Algebraic Equations using dummy derivatives," *Siam J. Sci. Comput.*, vol. 14, no. 3, pp. 677–692, 1993.

APPENDIX

A. Collecting proofs

1) Proof of Lemma 1: We consider the linear equation $Av{=}y$, where $y{\in}\mathbb{R}^n$ and v is the $\mathbb{R}^m$-valued unknown with coordinates $v_1,\ldots,v_m$. We assume $m{\geq}n$ otherwise the lemma is trivial.

We first prove the if part. We want to prove that, with the condition of the lemma, for every y, there exists a v such that $Av{=}y$ holds, almost everywhere when the non zero coefficients of A vary over some neighborhood. With a suitable renumbering of the coordinates of v, we can assume that Q is the identity matrix. Let P be a permutation matrix such that $A' =_{\text{def}} PA = \left[\begin{array}{cc} B_1 & B_2 \end{array}\right]$ where B_1 has a nonzero diagonal, and let σ be the permutation of $\{1,\ldots,n\}$ associated with P. Since $a'_{nn}{\neq}0$ we can express v_n in terms of $v_1\ldots v_{n-1}, y_{\sigma(n)}$ and then remove the last equation. This yields a reduced equation $A''v''{=}y''$, where A'' is $(n-1){\times}(m-1)$, $y''{\in}\mathbb{R}^{n-1}$, and v'' is the $\mathbb{R}^{m-1}$-valued unknown collecting $v_1\ldots v_{n-1}, v_{n+1},\ldots,v_m$. Matrix A'' has diagonal entries equal to $a''_{kk}{=}a'_{kk}-a'_{kn}/a'_{nn}$. Since a'_{kk} and a'_{nn} are non zero, the set of entries a'_{ij} of matrix A' causing $a''_{kk}{=}0$ for some k is exceptional in $\mathbb{R}^{n\times m}$ since it requires the condition $a'_{kn} = a'_{kk}a'_{nn}$ to hold. The corresponding set of entries a_{ij} of matrix $A = P^{-1}A'$ is exceptional as well, we denote by Ξ_n this exceptional subset of $\mathbb{R}^{n\times m}$. Thus we assume that the entries a_{ij} of matrix A do not belong to exceptional set Ξ_n. Since the diagonal entries of matrix A'' are all non zero, we can express v_{n-1} in terms of $v_1\ldots v_{n-2}, y_{\sigma(n-1)}$ and we are left with a further reduced equation $A'''v''' = y'''$ for which the same argument as before applies, namely: if the entries a_{ij} of matrix A do not belong to exceptional set Ξ_{n-1}, then the diagonal entries of A''' are all non zero. And so on, by reducing the number of equations and unknowns. We have thus proved the existence of exceptional sets $\Xi_n,\Xi_{n-1},\ldots,\Xi_1$ of $\mathbb{R}^{n\times m}$ such that, if the entries a_{ij} of matrix A do not belong to set $\Xi_n{\cup}\Xi_{n-1}{\cup}\cdots{\cup}\Xi_1$, the value of $v_n, v_{n-1},\ldots,v_1$ is uniquely determined as a function of the coordinates of y, by pivoting. This concludes the if part, since set $\Xi_n{\cup}\Xi_{n-1}{\cup}\cdots{\cup}\Xi_1$ is exceptional as well.

For the only-if part, assume that A is structurally onto, meaning that A is almost everywhere onto when its non-zero entries vary over some neighborhood. Call $V(A)$ the resulting set of neighbor matrices of A and let Ξ_n be the exceptional set for the entries $\tilde{a}_{ij}$ of $\tilde{A}$ outside which $\tilde{A}$ is onto, for $\tilde{A}$ ranging over $V(A)$. For each onto $\tilde{A}$, there exists a permutation matrix Q such that

$$\tilde{A}Q \;=\; \left[\begin{array}{cc} B_1 & B_2 \end{array}\right] \tag{138}$$

where B_1 is square invertible. Neighborhood $V(A)$ decomposes as the union $V(A) = \bigcup_\sigma V(A,\sigma)$, where $V(A,\sigma)$ collects the matrices $\tilde{A}$ for which decomposition (138) holds with the matrix Q_σ representing permutation σ. Since the number of permutations is finite, at least one of the $V(A,\sigma)$ has non-empty interior. Hence, without loss of generality we can assume that decomposition (138) holds with the same matrix Q. The map $V(A)\ni\tilde{A}\to\det(B_1)$ takes the value 0 on an exceptional set only. Now, we have $\det(B_1) = \sum_{i=1}^n \tilde{b}_{i1}\det(B_{i1})$, where B_{ij} is the submatrix of B_1 obtained by erasing row i and column j. Since B_1 is structurally nonsingular, there must be some $i\in\{1,\ldots,n\}$ such that $b_{i1}{\neq}0$ and B_{i1} is structurally nonsingular—otherwise we would have $\det(B_1)=0$ for any $\tilde{A}{\in}W(A)$, where $W(A)$ is some neighborhood contained in $V(A)$. Let P^1 be the permutation matrix exchanging rows 1 and i, so we replace $B_1 = B_1^0$ by $B_1^1 = P^1B_1^0$ and it is enough to prove the lemma for B_1^1. Now, since we have $b_{11}^1 = b_{i1}^1{\neq}0$ and B_{11}^1 is structurally nonsingular, we apply the same reasoning to B_{11}^1, which yields a second permutation matrix P^2 and so on. The wanted left permutation matrix is $P = P^m\ldots P^2P^1$.

2) Proof of Lemma 2: Since we reuse the same techniques as for the proof of Lemma 1, some details will be omitted.

We begin with the if part. Using (21), equation $Av{+}Cw{+}x{=}0$ becomes

$$0 \;=\; A_1v + C_1w + P_1x \tag{139}$$
$$0 \;=\; A_2v + P_2x \tag{140}$$
$$0 \;=\; A_3v + P_3x \tag{141}$$

where

$$P = \left[\begin{array}{c} P_1 \\ P_2 \\ P_3 \end{array}\right]$$

Eqn. (140) uniquely determines almost everywhere v as a partial function of x by using the same pivoting argument as for the proof of the if part of Lemma 1. Since v is fixed by (140), Eqn. (141) characterizes the consistent x. Eliminating w is performed by discarding eqn. (139). This proves the if part.

For the only if part, consider equation $Av + Cw + x = 0$ and use pivoting to eliminate w. To this end, pick an equation, say the i^{th} one, and a component of w, say the j^{th} one, such that w_j has a nonzero coefficient in the i^{th} equation. Let P_1'' be the permutation matrix exchanging the equations 1 and i, and let Q_1'' be the permutation matrix exchanging the rows 1 and j in w. Then, discard the first equation and the first coordinate of the permuted w and repeat the same argument on the remaining equations and components of the permuted w. Performing pivoting as above yields a second pair (P_2'',Q_2''), leaving invariant the first equation and the first component of w. The procedure repeats until step $r{\leq}q$, where no more component of w remains that has a nonzero coefficient in some equation. Setting

$$P_1' =_{\text{def}} P_r''\ldots P_1'' \text{ and } Q =_{\text{def}} Q_r''\ldots Q_1'' \tag{142}$$

yields

$$P_1'CQ = \left[\begin{array}{c} C_1' \\ 0 \end{array}\right]$$

where $r{\times}q$-matrix C_1' has rank r and allows using the r first permuted equations to eliminate w entirely. Accordingly,

$$C_1 =_{\text{def}} C_1'Q^{-1} \text{ is onto.} \tag{143}$$

Write

$$P_1'A = \begin{bmatrix} A_1' \\ A_2' \end{bmatrix} \text{ and } P_1' = \begin{bmatrix} P_{11}' \\ P_{12}' \end{bmatrix}$$

so that equation $Av + Cw + x = 0$ rewrites

$$0 = A_1'v + C_1w + P_{11}'x \quad (144)$$
$$0 = A_2'v + P_{12}'x \quad (145)$$

Eliminating w in equation $Av + Cw + x = 0$ amounts to discarding (144) and keeping only (145).

Now, let the nonzero coefficients of A and C vary over some neighborhood. The same reasoning used in the proof of the if part of Lemma 1 shows that

> a same matrix P_1' can work almost everywhere when the nonzero coefficients of A and C vary over some neighborhood. (146)

Thus we can now eliminate w and focus on equation (145) in which P_{11}' and P_{12}' are fixed. Apply to (145) the same pivoting argument we used to eliminate w in equation $Av+Cw+x = 0$ yields a permutation matrix P_2' such that

$$P_2'A_2' = \begin{bmatrix} A_2'' \\ 0 \end{bmatrix} \text{ where } A_2'' \text{ is structurally onto.} \quad (147)$$

Writing

$$P_2' = \begin{bmatrix} P_{21}' \\ P_{22}' \end{bmatrix}$$

equation (145) rewrites

$$0 = A_2''v + P_{21}'P_{12}'x \quad (148)$$
$$0 = P_{22}'P_{12}'x \quad (149)$$

Equation (149) characterizes the consistent x, whereas equation (148) is used to determine v, as a function of x.

Now, since A_2'' was obtained by pre- and post-multiplying A by permutation matrices and then taking a submatrix, the zero coefficients of A are preserved in A_2'' (modulo permutation and extraction) when the nonzero coefficients of A vary over some neighborhood. Hence (A, C) structurally nonsingular implies that A_2'' is structurally nonsingular too. Hence, by Corollary 1, there exists a permutation matrix $\widehat{P}_2$ such that $\widehat{P}_2A_2''$ has nonzero diagonal. Using (142), (146), and (147), and setting

$$P = \begin{bmatrix} I & 0 \\ 0 & \widetilde{P}_2 \end{bmatrix} P_1' \text{ , where } \widetilde{P}_2 = \begin{bmatrix} \widehat{P}_2 & 0 \\ 0 & I \end{bmatrix} P_2'$$

yields the desired permutation matrix P. This proves the only if part and the lemma is proved.

3) Proof of Theorem 1: We first prove the only if part. Assume that S possesses a causality analysis. Chose a total order on $\mathcal{Z} \uplus S$ that is an extension of order $\preceq_S$. This total order induces a renumbering of the equations E_i, thus defining a permutation matrix P, of dimension n, the number of equations in S. Using conditions 5) and 6) of Definition 4, we can rewrite equation $Av+Cw+x=0$ as follows:

$$\begin{array}{rl} \text{by 1) and 6):} & 0 = A_1v + [\, C_{11} \;\; C_{12} \,] \begin{bmatrix} w' \\ w'' \end{bmatrix} + P_1x \\ \text{by 5):} & 0 = A_2v + P_2x \\ \text{by 5):} & 0 = P_3x \end{array} \quad (150)$$

where C_{11} is structurally nonsingular, w' collects the variables of $\mathcal{W}_S \cap \mathcal{Z}$, possibly reordered, and permutation matrix P decomposes as

$$P = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix}$$

Setting $C_1 = \begin{bmatrix} C_{11} & C_{12} \end{bmatrix}$ shows that (150) is a particular form of (21), hence the pair (A, C) is structurally nonsingular by Lemma 2.

We move to the if part. Assume that (A, C) is structurally nonsingular. The reader is referred to the proof of Lemma 2, from which the following argument is reused. Using (21), equation $Av+Cw+x=0$ can be given the form (139,140,141). Eqn. (140) uniquely determines v as a partial function of x by using the same pivoting argument as for the proof of the if part of Lemma 1. Since v is fixed by (140), Eqn. (141) characterizes the consistent x. Eliminating w is performed by discarding eqn. (139), since C_1 is structurally onto. By using Corollary 1, we get a causality analysis for S.

4) Proof of Theorem 5: Consider the standard hybrid DAE system (110) and let $(x_s, b_s), s = (t, k) \in \mathbb{S}$, be a solution for it in the sense of Definition 10. Then, let λ be the parameter representing its initial condition. We show that this solution is the standardisation of some solution of hybrid dAE (115). Two cases can occur, see condition 1) of Definition 10:

Case 1: $t \notin T(\lambda)$. Then, by condition 1) of Definition 10, we can assume that $s = (t, 0)$, which we identify with t. Let $i \in I$ be the active mode, such that $b_{i,t} = 1$ holds. We then have $F_i(x_t^-, x_t, \dot{x}_t) = 0$. Pick $\mathbb{T} \ni \tau = t$. We can thus assume $\tau = [t_n]$ for some sequence t_n of reals converging to t. Hence,

$$\exists N \in \mathbb{N} \quad : \quad n \geq N \implies t_n \notin T(\lambda), \quad (151)$$

thus, $b_{i,(t_n,0)} = 1$. Therefore, setting $x_\tau = [x_{t_n}]$ and $\dot{x}_\tau = [\dot{x}_{t_n}]$, we get $b_i(x_\tau^-, x_\tau, \dot{x}_\tau) = [b_{i,t_n}] = 1$. Similarly, since F_i is smooth and the solution of $F_i(x^-, x, \dot{x}) = 0$ is assumed to be infinitely differentiable, we get $F_i(x_{t_n}^-, x_{t_n}, \dot{x}_{t_n}) = 0$ for $n \geq N$ where N is as in (151), whence $F_i(x_\tau^-, x_\tau, x_\tau') \approx 0$ follows as well. For this case, we thus proved the existence of a solution (x_τ, b_τ) for (115) such that $(x_t, b_t) = st(x_\tau, b_\tau)$.

Case 2: $t \in T(\lambda)$, meaning that one or more successive mode changes occur at t, so that the super-dense instants for consideration are $(t, 0), (t, 1), \ldots, (t, n_t(\lambda))$, on which the active mode changes at every instant. Suppose that $b_{i,(t,0)} = 1$, whence $x_{(t,0)}$ is a consistent value for $F_i = 0$. By condition 2 of Definition 10, we also have $b_{i,(t_n,0)} = 1$ for t_n any increasing sequence converging to t and n large enough. Consider $\mathbb{T} \ni \tau = [t_n]$, we have $st(\tau) = t$. Setting $x_\tau =_{\text{def}} [x_{(t_n,0)}]$ yields a consistent value for $F_i(x^-, x, x') = 0$ and complementing it with $b_i(x_\tau^-, x_\tau, x_\tau') = 1$ extends the solution of (115) at the considered τ. We further extend this solution for the subsequent non-standard instants $\tau + \partial, \ldots, \tau + n_t\partial$ as follows. First, observe that $\tau + \partial \approx \cdots \approx \tau + n_t\partial \approx t$. Then, we simply extend the solution of (115) by setting $b_{i,\tau+k\partial} = b_{i,(t,k)}$ and $x_{\tau+k\partial} = x_{(t,k)}$, where $(x_{(t,k)}, b_{i,(t,k)})$ is the given (standard) solution of (110). This proves the theorem.

B. A primer on non-standard analysis

The background material of this section is used in proofs, so the reader can skip it for a first reading. The text is borrowed verbatim from [4].

1) Intuitive introduction: We begin with an intuitive introduction to the construction of the non-standard reals. The goal is to augment $\mathbb{R}\cup\{\pm\infty\}$ by adding, to each x in the set, a set of elements that are "infinitesimally close" to it. We will call the resulting set ${}^\star\mathbb{R}$. Another requirement is that all operations and relations defined on $\mathbb{R}$ should extend to ${}^\star\mathbb{R}$.

A first idea is to represent such additional numbers as convergent sequences of reals. For example, elements infinitesimally close to the real number zero are the sequences $u_n = 1/n$, $v_n = 1/\sqrt{n}$ and $w_n = 1/n^2$. Observe that the above three sequences can be ordered: $v_n > u_n > w_n > 0$ where 0 denotes the constant zero sequence. Of course, infinitely large elements (close to $+\infty$) can also be considered, e.g., sequences $x_u = n$, $y_n = \sqrt{n}$, and $z_n = n^2$.

Unfortunately, this way of defining ${}^\star\mathbb{R}$ does not yield a total order since two sequences converging to zero cannot always be compared: if u_n and u'_n are two such sequences, the three sets $\{n \mid u_n > u'_n\}$, $\{n \mid u_n = u'_n\}$, and $\{n \mid u_n < u'_n\}$ may even all be infinitely large. The beautiful idea of Lindstrøm is to enforce that *exactly one of the above sets is important and the other two can be neglected*. This is achieved by fixing once and for all a finitely additive positive measure μ over the set $\mathbb{N}$ of integers with the following properties:[7]

1) $\mu : 2^{\mathbb{N}} \to \{0, 1\}$;
2) $\mu(X) = 0$ whenever X is finite;
3) $\mu(\mathbb{N}) = 1$.

Now, once μ is fixed, one can compare any two sequences: for the above case, exactly one of the three sets must have μ-measure 1 and the others must have μ-measure 0. Thus, say that $u > u', u = u'$, or $u < u'$, if $\mu(\{n \mid u_n > u'_n\} = 1)$, $\mu(\{n \mid u_n = u'_n\}) = 1$, or $\mu(\{n \mid u_n < u'_n\}) = 1$, respectively. Indeed, the same trick works for many other relations and operations on non-standard real numbers, as we shall see. We now proceed with a more formal presentation.

2) Non-standard domains: For I an arbitrary set, a *filter* $\mathcal{F}$ over I is a family of subsets of I such that:

1) the empty set does not belong to $\mathcal{F}$,
2) $P, Q \in \mathcal{F}$ implies $P \cap Q \in \mathcal{F}$, and
3) $P \in \mathcal{F}$ and $P \subset Q \subseteq I$ implies $Q \in \mathcal{F}$.

Consequently, $\mathcal{F}$ cannot contain both a set P and its complement P^c. A filter that contains one of the two for any subset $P \subseteq I$ is called an *ultra-filter*. At this point we recall Zorn's lemma, known to be equivalent to the axiom of choice:

Lemma 4 (Zorn's lemma): Any partially ordered set $(X, \leq)$ such that any chain in X possesses an upper bound has a maximal element.

A filter $\mathcal{F}$ over I is an ultra-filter if and only if it is maximal with respect to set inclusion. By Zorn's lemma, any filter $\mathcal{F}$ over I can be extended to an ultra-filter over I. Now, if I is infinite, the family of sets $\mathcal{F} = \{P \subseteq I \mid P^c \text{ is finite}\}$ is a *free* filter, meaning it contains no finite set. It can thus be extended to a free ultra-filter over I:

Lemma 5: Any infinite set has a free ultra-filter.

Every free ultra-filter $\mathcal{F}$ over I uniquely defines, by setting $\mu(P) = 1$ if $P \in \mathcal{F}$ and otherwise 0, a finitely additive measure[8] $\mu : 2^I \mapsto \{0, 1\}$, which satisfies

$$\mu(I) = 1 \text{ and, if } P \text{ is finite, then } \mu(P) = 0.$$

Now, fix an infinite set I and a finitely additive measure μ over I as above. Let $\mathbf{X}$ be a set and consider the Cartesian product $\mathbf{X}^I = (x_i)_{i\in I}$. Define $(x_i) \approx (x'_i)$ if and only if $\mu\{i \in I \mid x_i \neq x'_i\} = 0$. Relation $\approx$ is an equivalence relation whose equivalence classes are denoted by $[x_i]$ and we define

$$^\star\mathbf{X} = \mathbf{X}^I / \approx \tag{152}$$

$\mathbf{X}$ is naturally embedded into ${}^\star\mathbf{X}$ by mapping every $x \in \mathbf{X}$ to the constant tuple such that $x_i = x$ for every $i \in I$. Any algebraic structure over $\mathbf{X}$ (group, ring, field) carries over to ${}^\star\mathbf{X}$ by almost point-wise extension. In particular, if $[x_i] \neq 0$, meaning that $\mu\{i \mid x_i = 0\} = 0$ we can define its inverse $[x_i]^{-1}$ by taking $y_i = x_i^{-1}$ if $x_i \neq 0$ and $y_i = 0$ otherwise. This construction yields $\mu\{i \mid y_i x_i = 1\} = 1$, whence $[y_i][x_i] = 1$ in ${}^\star\mathbf{X}$. The existence of an inverse for any non-zero element of a ring is indeed stated by the formula: $\forall x\,(x = 0 \vee \exists y\,(xy = 1))$. More generally:

Lemma 6 (Transfer Principle): Every first order formula is true over ${}^\star\mathbf{X}$ if and only if it is true over $\mathbf{X}$.

3) Non-standard reals and integers: The above general construction can simply be applied to $\mathbf{X} = \mathbb{R}$ and $I = \mathbb{N}$. The result is denoted ${}^\star\mathbb{R}$; it is a field according to the transfer principle. By the same principle, ${}^\star\mathbb{R}$ is totally ordered by $[u_n] \leq [v_n]$ iff $\mu\{n \mid v_n > u_n\} = 0$. We claim that, for any finite $[x_n] \in {}^\star\mathbb{R}$, there exists a unique $st([x_n])$, call it the *standard part* of $[x_n]$, such that

$$st([x_n]) \in \mathbb{R} \quad \text{and} \quad st([x_n]) \approx [x_n]\,. \tag{153}$$

To prove this, let $x = \sup\{u \in \mathbb{R} \mid [u] \leq [x_n]\}$, where $[u]$ denotes the constant sequence equal to u. Since $[x_n]$ is finite, x exists and we only need to show that $[x_n] - x$ is infinitesimal. If not, then there exists $y \in \mathbb{R}, y > 0$ such that $y < |x - [x_n]|$, that is, either $x < [x_n] - [y]$ or $x > [x_n] + [y]$, which both contradict the definition of x. The uniqueness of x is clear, thus we can define $st([x_n]) = x$. Infinite non-standard reals have no standard part in $\mathbb{R}$.

It is also of interest to apply the general construction (152) to $\mathbf{X} = I = \mathbb{N}$, which results in the set ${}^\star\mathbb{N}$ of *non-standard natural numbers*. The non-standard set ${}^\star\mathbb{N}$ differs from $\mathbb{N}$ by the addition of *infinite natural numbers,* which are equivalence classes of sequences of integers whose essential limit is $+\infty$.

[7] The existence of such a measure is non trivial and is explained later.

[8] Observe that, as a consequence, μ cannot be sigma-additive (in contrast to probability measures or Radon measures) in that it is *not* true that $\mu(\bigcup_n A_n) = \sum_n \mu(A_n)$ holds for an infinite denumerable sequence A_n of pairwise disjoint subsets of $\mathbb{N}$.

4) Integrals and ODE: Any sequence (g_n) of functions $g_n : \mathbb{R} \mapsto \mathbb{R}$ point-wise defines a function $[g_n] : {}^\star\mathbb{R} \mapsto {}^\star\mathbb{R}$ by setting

$$[g_n]([x_n]) = [g_n(x_n)] \tag{154}$$

A function ${}^\star\mathbb{R} \to {}^\star\mathbb{R}$ so obtained is called *internal*. Properties of and operations on ordinary functions extend point-wise to internal functions of ${}^\star\mathbb{R} \to {}^\star\mathbb{R}$. The *non-standard version* of $g : \mathbb{R} \to \mathbb{R}$ is the internal function ${}^\star g = [g, g, g, \dots]$. The same notions apply to sets. An internal set $A = [A_n]$ is called *hyperfinite* if $\mu\{n \mid A_n \text{ finite}\} = 1$; the *cardinal* $|A|$ of A is defined as $[|A_n|]$.

Now, consider an infinite number $N \in {}^\star\mathbb{N}$ and the set

$$T = \left\{ 0, \frac{1}{N}, \frac{2}{N}, \frac{3}{N}, \dots \frac{N-1}{N}, 1 \right\} \tag{155}$$

By definition, if $N = [N_n]$, then $T = [T_n]$ with

$$T_n = \left\{ 0, \frac{1}{N_n}, \frac{2}{N_n}, \frac{3}{N_n}, \dots \frac{N_n - 1}{N_n}, 1 \right\}$$

hence $|T| = [|T_n|] = [N_n + 1] = N + 1$. Now, consider an internal function $g = [g_n]$ and a hyperfinite set $A = [A_n]$. The *sum* of g over A can be defined:

$$\sum_{a \in A} g(a) =_{\text{def}} \left[\sum_{a \in A_n} g_n(a) \right]$$

If t is as above, and $f : \mathbb{R} \to \mathbb{R}$ is a standard function, we obtain

$$\sum_{t \in T} \frac{1}{N} {}^\star f(t) = \left[\sum_{t \in T_n} \frac{1}{N_n} f(t_n) \right] \tag{156}$$

Now, f continuous implies $\sum_{t \in T_n} \frac{1}{N_n} f(t_n) \to \int_0^1 f(t)dt$, so,

$$\int_0^1 f(t)dt = st\left(\sum_{t \in T} \frac{1}{N} {}^\star f(t) \right) \tag{157}$$

Under the same assumptions, for any $t \in [0, 1]$,

$$\int_0^t f(u)du = st\left(\sum_{u \in T, u \le t} \frac{1}{N} {}^\star f(t) \right) \tag{158}$$

Now, consider the following ODE:

$$\dot{x} = f(x, t), \quad x(0) = x_0 \tag{159}$$

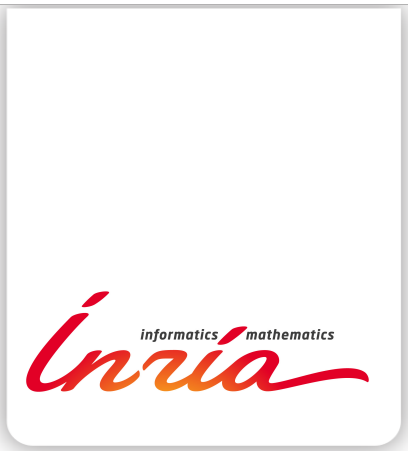
Assume (159) possesses a solution $[0, 1] \ni t \mapsto x(t)$ such that the function $t \mapsto f(x(t), t)$ is continuous. Rewriting (159) in its equivalent integral form $x(t) = x_0 + \int_0^t f(x(u), u)du$ and using (158) yields

$$x(t) = st\left(x_0 + \sum_{u \in T, u \le t} \frac{1}{N} {}^\star f(x(u), u) \right) \tag{160}$$

The substitution in (160) of $\partial = 1/N$, which is positive and infinitesimal, yields $T = \{t_n = n\partial \mid n = 0, \dots, N\}$. The expression in parentheses on the right hand side of (160) is the piecewise-constant right-continuous function ${}^\star x(t), t \in [0, 1]$ such that, for $n = 1, \dots, N$:

$$\begin{array}{rcl} {}^\star x(t_n) & = & {}^\star x(t_{n-1}) + \partial \times {}^\star f({}^\star x(t_{n-1}), t_{n-1}) \\ {}^\star x(t_0) & = & x_0 \end{array} \tag{161}$$

By (160), the solutions x, of ODE (159), and ${}^\star x$, as computed by algorithm (161), are related by $x = st({}^\star x)$. Formula (161) can be seen as a *non-standard semantics* for ODE (159); one which depends on the choice of infinitesimal step parameter ∂. Property (160), though, expresses the idea that all these non-standard semantics are equivalent from the standard viewpoint regardless of the choice made for ∂. This fact is referred to as the *standardization principle*.



RESEARCH CENTRE
RENNES – BRETAGNE ATLANTIQUE

Campus universitaire de Beaulieu
35042 Rennes Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399